IDENTIFICATION OF DIFFERENTIALLY METHYLATED
AND MUTATED NUCLEIC ACIDS

RELATED U.S. APPLICATION DATA:

5   This application is a continuation-in-part of U.S. Ser. No. 09/163,951, filed Sep. 30, 1998, which is a continuation-in-part of U.S. Ser. No. 08/657,866, filed May 31, 1996, now U.S. Pat. No. 5,871,917 incorporated by reference herein, the benefit of whose filing date is claimed pursuant to 35 U.S.C. § 120.

10   FIELD OF THE INVENTION:

The present invention provides Methyl- (or Mutant-) Differential Display (MDD) methods and nucleic acid probes for detecting mutations and/or the methylation patterns of nucleic acids. Genes are frequently not methylated in the cells where they are expressed but are methylated in cell types where they are not expressed. Moreover, tumor cell DNA is

15   frequently methylated to a different extent and in different regions than is the DNA of normal cells. The present invention is used for identifying which regions of the genome are methylated or mutated in different cell types, including cancerous cell types. The present invention also is used for identifying whether a tissue sample has a methylation or mutation pattern which is normal or like that of known cancer cells. The present invention can be

20   used to identify aberrant expression of known or unknown genes. Such aberrant expression may occur at an incorrect location or an incorrect time, i.e. tissue or developmental specificity is lost. Genes and genomic regions discovered according to the present invention are useful for identifying products, e.g. mRNA, protein or fragments thereof, which are aberrantly expressed.

25

BACKGROUND OF THE INVENTION:

DNA is often methylated in normal mammalian cells. For example, DNA is methylated to determine whether a given gene will be expressed and whether the maternal or the paternal allele of that gene will be expressed. (Little and Wainwright, 1995). While

30   methylation is known to occur at CpG sequences, only recent studies indicate that CpNpG

methylation is known to occur at CpG sequences, only recent studies indicate that CpNpG sequences may be methylated (Clark et al., 1995). Methylation at CpG sites has been much more widely studied and is better understood.

Methylation occurs by enzymatic recognition of CpG and CpNpG sequences followed by placement of a methyl (CH_3) group on the fifth carbon atom of a cytosine base. The enzyme that mediates methylation of CpG dinucleotides, 5-cytosine methyltransferase, is essential for embryonic development -- without it embryos die soon after gastrulation. It is not yet clear whether this enzyme methylates CpNpG sites (Laird and Jaenisch, 1994).

When a gene has many methylated cytosines it is less likely to be expressed (Willson, 1991). Hence, if a maternally-inherited gene is more highly methylated than the paternally-inherited gene, the paternally-inherited gene will generally give rise to more gene product. Similarly, when a gene is expressed in a tissue-specific manner, that gene will often be unmethylated in the tissues where it is active, but will be highly methylated in the tissues where it is inactive. Incorrect methylation is thought to be the cause of some diseases including Beckwith-Wiedemann syndrome and Prader-Willi syndrome (Henry et al., 1991; Nicholls et al., 1989).

The methylation patterns of DNA from tumor cells are generally different than those of normal cells (Laird, 1994). Tumor cell DNA is generally undermethylated relative to normal cell DNA, but selected regions of the tumor cell genome may be more highly methylated than the same regions of a normal cell's genome. Hence, detection of altered methylation patterns in the DNA of a tissue sample is an indication that the tissue is cancerous. For example, the gene for Insulin-Like Growth Factor 2 (IGF2) is hypomethylated in a number of cancerous tissues, such as Wilm's Tumors, rhabdomyosarcoma, lung cancer and hepatoblastomas (Rainier et al., 1993; Ogawa et al., 1993; Zhan et al., 1994; Pedone et al., 1994; Suzuki et al., 1994; Rainier et al., 1995).

The present invention is directed to a method of detecting differential methylation at CpNpG sequences by cutting test and control DNAs with a restriction enzyme that will not cut methylated DNA, and then detecting the difference in size of the resulting restriction fragments.

While methylation-sensitive restriction enzymes have been used for observing differential methylation in various cells, no commercial assays exist for use on human samples because differentially methylated sequences represent such a minute proportion of the human genome that they are not readily detected. The human genome is both highly

5 complex, in that it contains a great diversity of DNA sequences, and highly repetitive, in that it contains a lot of DNA with very similar or identical sequences. The high complexity and repetitiveness of human DNA confounds efforts at detecting and isolating the minute amount of differentially methylated DNA which may be present in a test sample. The present invention remedies this detection problem by providing new procedures for screening a

10 selected subset of the mammalian genome which is most likely to contain genetic functions.

The present invention provides techniques for detecting and isolating differentially methylated or mutated segments of DNA which may be present in a tissue sample in only minute amounts by using one or more rounds of DNA amplification coupled with subtractive hybridization to identify such segments of DNA. DNA amplification has been

15 coupled with subtractive hybridization in the Representational Difference Analysis (RDA) procedures disclosed in U.S. Patent 5,436,142 to Wigler et al. and Lisitsyn et al. (1993). However, for the subtractive hybridization step of such RDA procedures to proceed in a reasonable time and with reasonable efficiency, only a subset of the genome can be examined. To accomplish this necessary reduction in the complexity of the sample DNA,

20 Wigler et al. and Lisitsyn et al. disclose cutting DNA samples with restriction enzymes that cut infrequently and randomly. However, selection of enzymes which randomly cut the genome means that the portion of the genome which is examined is not enriched for any particular population of DNA fragments. Thus, when RDA is used, only a random subset of the human genome, which includes repetitive elements, noncoding regions and other

25 sequences which are generally not of interest, can be tested in a single experiment.

In contrast, the present invention is directed to methods which use enzymes that cut frequently and that specifically cut CG-rich regions of the genome. These enzymes are chosen because CG-rich regions of the genome are not evenly distributed in the genome -- instead, CG-rich regions are frequently found near genes, and particularly near the promoter

30 regions of genes. This means that the proportion of the genome that is examined by the

present methods will be enriched for genetically-encoded sequences as well as for regulatory sequences. Moreover, unlike the RDA method, the present methods selectively identify regions of the genome which are hypomethylated or hypermethylated by using enzymes which specifically cut non-methylated CG-rich sequences. The present invention therefore represents an improvement over RDA methods because of its ability to select DNA fragments which are likely to be near or to encode genetic functions.

SUMMARY OF THE INVENTION:

The present invention provides probes and methods of detecting whether a CNG triplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cell.

The present invention provides a method of detecting whether a CNG triplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cells which includes:

a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;

b) cleaving the control-cell DNA and the test-cell DNA with a master restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;

c) preparing a probe from a DNA isolated by the present methods, for example, a DNA selected from the group consisting of SEQ ID NO:7-10;

d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and

e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex;

wherein the master restriction enzyme cleaves a nonmethylated CNG DNA sequence but does not cleave a methylated CNG DNA sequence.

Similarly, the present invention provides a method of detecting whether a DNA site is mutated in a genomic DNA present in a test sample of cells which includes steps a) through e) above but where a detector restriction enzyme is used instead of the master

restriction enzyme, wherein the detector restriction enzyme does not cleave a mutated DNA site but does cleave a corresponding nonmutated DNA site.

The present invention also provides methods of isolating probes to detect hypomethylation or hypermethylation in a CNG triplet of DNA. In particular, the present invention provides a method of isolating a probe to detect hypomethylation in a CNG triplet of DNA which includes:

a) cleaving a tester sample of genomic DNA with both a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

b) ligating a first set of adaptors onto master enzyme cut DNA ends of the cleaved tester sample to generate a first-tester amplification template;

c) amplifying the first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors;

d) cleaving off the first adaptors from the first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) melting and hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product;

f) adding nucleotides onto DNA ends present in the mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

f) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect the hypomethylation or the hypermethylation in a CNG triplet of DNA;

wherein the master restriction enzyme cleaves a nonmethylated CNG DNA sequence but does not cleave a methylated CNG DNA sequence;

wherein the partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein the driver DNA is.cut with both the master restriction enzyme and the partner restriction enzyme and amplified using primers that recognize DNA ends cut by the master restriction enzyme.

5 Similarly, when the present invention provides a method of isolating a probe to detect hypermethylation in a CNG triplet of DNA, normal DNA is used instead of the tester DNA and the normal amplicon is annealed and hybridized with a molar excess of tester driver DNA.

Moreover the present methods are used to isolate probes to detect a mutation in a test sample of genomic DNA by using a detector enzyme instead of a partner restriction

10 enzyme, wherein the detector restriction enzyme cleaves a normal DNA site but not a mutant DNA site.

The present invention further provides a kit for detecting hypomethylation in a CNG triplet of DNA which is present in a test tissue sample which includes a DNA or a probe isolated by the methods of the present invention, for example, a DNA having SEQ ID

15 NO:7-10.

The present invention provides nucleic acid sequences which are differentially methylated in human breast and ovarian cancer tissues, as compared to normal tissue.

The present invention provides nucleic acid sequences which are useful as probes for determining the methylation state of genomic DNA sequences which are differentially

20 methylated in tumor tissues.

Similarly, the present invention provides nucleic acid sequences which are useful as probes for the detection of mRNA expression, as well as oligonucleotide sequences which can be used to amplify and thus detect normal and aberrant expression of particular mRNA species.

25 The DNA sequences provided herein may also be expressed in host cells for the production of the encoded protein or protein fragments.

The proteins and protein fragments expressed from the nucleotide sequences disclosed herein are useful for the production of monoclonal antibodies (Mabs). Such Mabs have a variety of uses including, but not limited to, detection of normal or aberrant

30 protein expression in human tissues and fluids.

Antibodies of the invention bind selectively to tsp50 protein wherein tsp50 protein comprises SEQ ID NO:16 from about amino acid residue number 1 to about amino acid residue number 385. The term "antibody" is used herein to include complete antibodies (e.g., bivalent IgG, pentavalent IgM) as well as fragments of antibodies which contain an antigen binding site. Such fragment include, but are not limited to Fab, F(ab')$_2$, Fv and single chain Fv (scFv) fragments. Such fragments may or may not include antibody constant domains. For example, Fab's lack constant domains which are required for complement fixation. scFvs are composed of an antibody variable light chain (V_L) linked to a variable heavy chain (V_H) by a flexible linker. scFvs are able to bind antigen and can be rapidly produced in bacteria. The invention includes antibodies and antibody fragments which are produced in bacteria and in mammalian cell culture. An antibody obtained from a bacteriophage library can be a complete antibody or an antibody fragment. Generally, the domains actually present in such a library are heavy chain variable domains (V_H) and light chain variable domains (V_L) which together comprise Fv or scFv, with the addition, in some cases, of a heavy chain constant domain (C_H1) and a light chain constant domain (C_L). The four domains (i.e., V_H-C_H1 and V_L-C_L) comprise an Fab. Complete antibodies are obtained from such a library by replacing missing constant domains once a desired V_H-V_L combination has been identified.

Antibodies of the invention can be monoclonal antibodies (Mab) or polyclonal antibodies. Antibodies of the invention which are useful for *in vitro* detection of tsp50 protein can be from any source, and in addition may be chimeric antibodies which comprise sequences of amino acids from two or more sources. Sources include but are not limited to a mouse or a rat or a human. Preferred antibodies for use *in vivo* (i.e., detection or treatment of cancer) have reduced antigenicity in humans, and more preferably are not antigenic in humans. Chimeric antibodies for use *in vivo* contain human amino acid sequences and include humanized antibodies which are non-human antibodies substituted with sequences of human origin to reduce or eliminate immunogenicity, but which retain the binding characteristics of the non-human antibody. Particularly preferred antibodies are human antibodies.

Polyclonal antibodies can be obtained directly from the serum of animals immunized with tsp50. Human polyclonal antibodies can be obtained from transgenic animals which comprise human Ig genes. Human antibodies can also be obtained from combinatorial libraries comprising cells expressing combinations of human heavy and light chains. Human antibodies have human heavy chains and human light chains. Human antibodies also include variants which have been created to, for example, improve antigen binding characteristics. In a normal immune response, this occurs through somatic mutation and selection. *In vitro*, this is accomplished by mutation of one or more residues, followed by screening for antibodies having preferred binding characteristics. Preferred binding characteristics are improved affinity and/or improved specificity.

Bacteriophage libraries used to obtain antibodies of the invention comprise combinations of rearranged Ig genes from mature mammalian B cells. Animals, including animals comprising human Ig genes, from which these libraries are developed may be first immunized with tsp50 to increase the proportion of the library which is reactive with tsp50. Human antibodies of the invention can therefore be obtained from bacteriophage libraries made from immunized transgenic animals comprising human Ig genes, which animals may first have been immunized with tsp50. Alternatively, human antibodies can be obtained from bacteriophage libraries developed directly from human immune system cells. Human antibodies further comprise those which have been obtained by these or similar methods which are then modified to improve their binding characteristics.

The invention also provides chimeric anti-tsp50 antibodies which have binding sites derived from non-human anti-tsp50 antibodies. To minimize immunogenicity in humans, chimeric anti-tsp50 antibodies of the invention comprise non-binding site amino acids which have been substituted with human sequences. The humanized regions of chimeric anti-tsp50 antibodies can be heavy and light chain constant domains. The humanized regions can also be framework regions of variable domains. The humanized regions can also be constant region and variable region amino acids which are exposed on the surface of the antibody molecule. Chimeric and humanized antibodies of the invention further comprise chimeric antibodies and humanized antibodies which have been modified to improve their binding

characteristics. Improved binding can mean greater affinity and/or greater specificity. Such modification can be the result of, for example, *in vitro* mutation and selection.

Antibodies of the invention may be linked to an anti-tumor agent or a detectable label. This allows the antibody to target the anti-tumor agent or

detectable label to the tumor. Thus, antibodies of the invention can be suitable for use in a method of treatment of the human body or in a method of diagnosis applied directly to the human body, or in a method of diagnosis applied to samples of human tissues or body fluids obtained from the human body. A review of the use of antibodies in diagnosis and therapy is provided by Waldmann (1991).

The invention includes pharmaceutical compositions which comprise the aforementioned antibodies and antibody conjugates.

BRIEF DESCRIPTIONS OF THE DRAWINGS:

Fig. 1 provides a schematic diagram of how the MDD technique generally selects for DNA fragments which have non-methylated CG-rich ends. Driver DNA is isolated from a normal cell or tissue sample of the same tissue-type as the tester cell DNA. Tester DNA is isolated from a cell or tissue sample which likely has a mutation or a difference in the extent to which its DNA is methylated. Tester and driver DNAs are cleaved by the master (M) and partner (P) restriction enzymes. The master enzyme will not cleave DNA at sites that have a methyl group (superscript "m") but will cleave at sites which are non-methylated (superscript m having a line drawn through it). The partner enzyme cleaves the DNA into a size and a complexity appropriate for DNA amplification and subtractive hybridization. After cleavage, a small target fragment is completely excised from the tester DNA but not from the driver DNA. Adaptors (m-linkers) which hybridize only to master-enzyme-cut DNA ends are ligated onto the tester and driver DNA fragments. Unique target DNAs from the tester sample will have adaptors on both ends. Upon DNA amplification using a primer that recognizes only the adaptor sequences, only DNA fragments with adaptors on both ends will be efficiently amplified.

Fig. 2 depicts an agarose gel containing electrophoretically separated DNA fragments isolated by MDD from malignant B-cells derived from Chronic Lymphocytic

Leukemia (CLL) patients. Lane A contains the different products isolated from CLL

patient #111. Lane M contains *Hae* III 0174 DNA size markers. As illustrated, the

present MDD methods give rise to discrete fragments of DNA which can readily be

isolated and subcloned or used as probes for identifying the methylation patterns of genomic

5   DNA samples from different cell types.

Fig. 3 depicts a Southern blot of *Msp* I-digested tester (T) and driver (D) DNA

hybridized with the *CLL58* probe. The lane labeled "probe #58" contains a clone of the

CLL58 probe. The lane labeled "T" contains amplified 2μg CD5$^+$ B-cell DNA, whereas

the lane "D" contains 2μg amplified neutrophil DNA. Both T and D DNA samples were

10   isolated from the CLL patient #111 and amplified by the MDD method. As a positive

control, probe *CLL58* hybridizes strongly to itself. Probe *CLL58* also hybridizes strongly

with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA

hypomethylation in the tester DNA.

Fig. 4 depicts a Southern blot of *Msp* I-digested DNA hybridized with the *CLL58*

15   probe. The lane labeled "probe #58" contains a clone of the *CLL58* probe. The lane

labeled "T" contains 6 μg of CD5$^+$ B-cell DNA, whereas lane "N" contains 6 μg of

neutrophil DNA, each isolated from the CLL patient #111. As a positive control, probe

CLL58 hybridizes strongly to itself. Probe *CLL58* hybridizes with a lower band and

several upper bands in B-cell DNA (lane T), while comparatively hybridizing much less with

20   the lower band in neutrophil cell DNA (lane N). The lesser hybridization to the lower band

in neutrophil DNA is due to DNA methylation which prevents complete excision of that

fragment. Hence, the *CLL58* probe provides a comparative methyl-differential display in

B-cells and neutrophil cells.

Fig. 5 depicts an agarose gel after electrophoresis of the different final products

25   isolated from breast patient #13 (lane A), patient #14 (lane B) and breast patient #4 (lane

C) using the MDD technique. Lane MW contains *Hae* III 0174 DNA size markers.

Fig. 6a depicts a Southern blot of *Msp* I-digested tester (T) and driver (D) DNA

hybridized with the *BR50* probe. The lane labeled "probe #50" contains a clone of the

BR50 probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the

30   lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples were

isolated from the breast patient #14 and amplified by the MDD method. As a positive control, probe *BR50* hybridizes strongly to itself. Probe *BR50* also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA hypomethylation in the tester DNA.

5   Fig. 6b depicts a Southern blot of *Msp* I-digested tester (T) and driver (D) DNA hybridized with the *BR104* probe. The lane labeled "probe #104" contains a clone of the *BR104* probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples were isolated from the breast patient #13 and amplified by the MDD method. As a positive

10   control, probe *BR104* hybridizes strongly to itself. Probe *BR104* also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA hypomethylation in the tester DNA.

Fig. 6c depicts a Southern blot of *Msp* I-digested tester (T) and driver (D) DNA hybridized with the *BR254* probe. The lane labeled "probe #254" contains a clone of the

15   *BR254* probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples were isolated from the breast patient #4 and amplified by the MDD method. As a positive control, probe *BR254* hybridizes strongly to itself. Probe *BR254* also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA

20   hypomethylation in the tester DNA.

Fig. 7a depicts a genomic Southern blot of *Msp* I-digested DNA hybridized with a probe *BR50* isolated from a breast cancer patient. The lane labeled "probe #50" contains a clone of the isolated *BR50* probe. The lane labeled "T" contains 6 μg of B cell DNA, whereas lane "N" contains 6 μg of normal breast DNA, each isolated from the breast

25   cancer patient #14. As a positive control, probe *BR50* hybridizes strongly to itself. Probe *BR50* hybridizes strongly with a lower band and weakly with a medium band in breast tumor DNA (lane T). However, in normal breast cell DNA (lane N), probe *BR50* hybridizes much less with the lower band and only weakly to an upper band. The lesser hybridization to the lower band in normal breast DNA is due to DNA methylation which

prevents complete excision of that fragment. Hence, the *BR50* probe provides a comparative methyl-differential display in tumor and normal breast cells.

Fig. 7b depicts a genomic Southern blot of *Msp* I-digested DNA hybridized with a probe *BR104* isolated from a breast cancer patient. The lane labeled "probe #104"

5   contains a clone of the isolated *BR104* probe. The lane labeled "T" contains 6 μg of tumor DNA, whereas lane "N" contains 6 μg of normal breast cell DNA, each isolated from the breast cancer patient #13. As a positive control, probe *BR104* hybridizes strongly to itself. Probe *BR104* hybridizes strongly with a lower band and weakly with a medium band in tumor DNA (lane T). However, in normal breast cell DNA (lane N), probe *BR104*

10   hybridizes lightly with the lower band and moderately with middle and upper bands. The lesser hybridization to the lower band in normal breast DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the *BR104* probe provides a comparative methyl-differential display in tumorous and normal breast cells.

Fig. 7c depicts a genomic Southern blot of *Msp* I-digested DNA hybridized with a

15   probe *BR254* isolated from a breast cancer patient. The lane labeled "probe #254" contains a clone of the isolated *BR254* probe. The lane labeled "T" contains 6 μg of B-cell DNA, whereas lane "N" contains 6 μg of normal breast cell DNA, each isolated from the breast cancer patient #4. As a positive control, probe *BR254* hybridizes strongly to itself. Probe *BR254* hybridizes strongly with a lower band and with an upper band in tumor DNA

20   (lane T). However, in normal breast cell DNA (lane N), probe *BR254* hybridizes only with the upper band. The lack of hybridization to the lower band in normal breast DNA is due to DNA methylation which prevents excision of that fragment. Hence, the *BR254* probe provides a comparative methyl-differential display in tumorous and normal breast cells.

Fig. 8 depicts a Southern blot of *Msp* I-digested genomic DNA from five different

25   breast cancer patients which was hybridized with the *BR50* probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #50" contains a clone of the isolated *BR50* probe. The lane labeled "T" contains 6 μg of DNA isolated from a tumor, whereas lane "N" contains 6 μg of DNA from normal breast

30   tissue. As a positive control, probe *BR50* hybridizes strongly to itself. For each of the five

patients, probe *BR50* hybridizes strongly with a lower band and less strongly with an upper band in the tumor DNA samples (T). However, probe *BR50* hybridizes much less with the lower band in the normal DNA samples (N). The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment.

5   Hence, the *BR50* probe provides a comparative methyl-differential display in breast tumor and normal DNA samples.

Fig. 9 depicts a Southern blot of *Msp* I-digested genomic DNA from four different breast cancer patients which was hybridized with the *BR104* probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched

10   normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #104" contains a clone of the isolated *BR104* probe. As a positive control, probe *BR104* hybridizes strongly to itself. The lanes labeled "T" contain 6 μg of DNA isolated from a breast tumor, whereas the lanes labeled "N" contain 6 μg of DNA from normal breast tissue. For tumor DNA from each of the four patients, probe *BR104* hybridizes with a

15   lower band and a middle band. In the normal DNA, the probe hybridizes mainly with the middle band and an upper band. The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the *BR104* probe provides a comparative methyl-differential display in breast tumor and normal breast DNA samples.

20   Fig. 10 depicts a Southern blot of *Msp* I-digested genomic DNA from four different breast cancer patients which was hybridized with the *BR254* probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #254" contains a clone of the isolated *BR254* probe. The lane labeled "T" contains 6 μg of

25   DNA isolated from a breast tumor, whereas lane "N" contains 6 μg of DNA from normal breast tissue. As a positive control, probe *BR254* hybridizes strongly to itself. For each of the four patients, probe *BR254* hybridizes strongly with a lower band and less strongly with an upper band in the tumor DNA samples (T). However, probe *BR254* hybridizes much less with the lower band in the normal DNA samples (N). The lesser hybridization to the

30   lower band in normal DNA is due to DNA methylation which prevents complete excision of

that fragment. Hence, the *BR254* probe provides a comparative methyl-differential display in breast tumor and normal DNA samples.

Fig. 11 depicts an amplification pattern detected by probe *BR254* in DNA isolated from a breast tumor biopsy. Fig. 11 provides a Southern blot of normal (N) and tumor (T)

5   DNA from breast cancer patient #23 which was digested with *Msp* I and hybridized with the *BR254* probe. Both the N and T lanes contain 6 μg of DNA. The lane labeled "probe #254" contains a clone of the isolated *BR254* probe, which hybridizes to itself. In the tumor DNA, the *BR254* probe hybridizes very, very strongly to a DNA fragment to which it hybridizes only moderately in normal DNA. This indicates that the breast tumor DNA of

10   patient #23 has a highly amplified region of genomic DNA which is not present in her normal DNA.

Fig. 12 depicts a Southern blot of *Msp* I-digested genomic DNA from five different ovarian cancer patients which was hybridized with the *BR50* probe. The number for each patient is listed above the bracket. The hybridization patterns of 6 μg primary tumor DNA

15   (T or T1), 6 μg metastatic tumor DNA (Tm) and 6 μg normal ovarian DNA (N) from these patients are provided for comparison. The lane labeled "probe #50" contains a clone of the isolated *BR50* probe. The lane labeled "T" contains 6 μg of DNA isolated from a ovarian tumor, whereas lane "N" contains 6 μg of DNA from normal ovarian tissue. As a positive control, probe *BR50* hybridizes strongly to itself. In primary tumor DNAs, the

20   probe predominantly hybridizes with the lower bands, and slightly hybridizes with the upper bands. In metastatic tumor DNAs, the probe only hybridizes with the lower bands. In normal DNAs, the probe hybridized with the lower bands and the upper bands. The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the *BR50* probe provides a comparative

25   methyl-differential display in ovarian tumor and normal DNA samples. The degree of DNA hypomethylation may be related to the progress of the disease.

Fig. 13 depicts a Southern blot of *Msp* I-digested genomic DNA from four different ovarian cancer patients which was hybridized with the *BR254* probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched

30   normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe

#254" contains a clone of the isolated *BR254* probe. The lane labeled "T" contains 6 μg of DNA isolated from a ovarian tumor, whereas lane "N" contains 6 μg of DNA from normal ovarian tissue. As a positive control, probe *BR254* hybridizes strongly to itself. For each of the four patients, probe *BR254* hybridizes with a lower band and with an upper band in the tumor DNA samples (T). However, probe *BR254* hybridizes only with the upper band in the normal DNA samples (N). The lack of hybridization to the lower band in normal DNA is due to DNA methylation which prevents excision of that fragment. Hence, the *BR254* probe provides a comparative methyl-differential display in ovarian tumor and normal DNA samples.

Fig. 14 depicts a Southern blot of *Msp* I-digested genomic DNA from six different colon cancer patients which was hybridized with the *BR254* probe. The number for each patient is listed above the bracket. The hybridization patterns of 6 μg tumor (T) and 6 μg matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #254" contains a clone of the isolated *BR254* probe as a positive control to show that probe *BR254* hybridizes strongly to itself. In the tumor DNA, the probe hybridizes with lower and upper bands with equal intensity. In the normal DNA, the probe hybridizes mainly with the lower band. Hence, the *BR254* probe detects alteration of DNA methylation in colon cancer DNA.

Fig. 15 depicts a Northern blot of mRNA samples from several normal human tissues, which was hybridized with probe *BR50*. For quantitative comparison, the same membrane was hybridized to a control probe which is expressed in all tissues. As indicated by the band labeled "TSP50", probe *BR50* hybridizes strongly with testes specific mRNA, but not with mRNA from other sources.

Fig. 16 depicts a Southern blot of DNA samples obtained from nine normal human tissues. 6 μg of DNA from each tissue was cleaved with *Hpa* II (panel a) or *Msp* I (panel b) and analyzed with probe *BR50*. Among DNA samples digested with *Hpa* II, hybridized digestion products (approximately 1 kb and 2 kb) are observed only for testes tissue. For DNA digested with *Msp* I, one or both products are observed in all samples. Hybridization of the control probe to a single band in each digested DNA sample indicates that digestion was complete.

Fig. 17 depicts differential expression of TSP50 in five breast tumor (T) and normal (N) tissues as determined by RT-PCR. Lane M contains *Hae* III digested φ174 molecular weight markers. Positive controls were generated using cDNA prepared from testes tissue and primers specific for TSP50 and β-actin.

DETAILED DESCRIPTION OF THE INVENTION:

The present invention provides methods and probes for detecting and isolating DNA sequences which are mutated or methylated in one tissue type but not in another. These methods are herein named Methyl- (or Mutant-) Differential Display ("MDD") methods. These MDD methods, and probes that are isolated by MDD, are useful *inter alia* for identifying genes that are expressed in a tissue-specific manner. Moreover, the present MDD methods and probes are useful for detecting cancerous cells, viruses, bacteria, point mutations, amplifications, deletions and genomic rearrangements.

To detect whether a DNA sequence is differentially methylated, the present invention uses methylation-sensitive restriction enzymes to cleave tester and control DNAs. These methylation-sensitive restriction enzymes will cut their DNA recognition sites when those sites are not methylated but do not cut the DNA site if it is methylated. Hence, a non-methylated tester DNA will be cut into smaller sizes than a methylated control DNA. Similarly, a hypermethylated tester DNA will not be cleaved and will give rise to larger fragments than a normally non-methylated control DNA.

The present invention is particularly useful for detecting and isolating DNA fragments that are normally methylated but which, for some reason, are non-methylated in a small proportion of cells. Such DNA fragments may normally be methylated for a number of reasons. For example, such DNA fragments may be normally methylated because they contain, or are associated with, genes that are rarely expressed, genes that are expressed only during early development, genes that are expressed in only certain cell-types, and the like.

While highly expressed genes are frequently unmethylated, those highly expressed genes are often not very interesting because they are "housekeeping genes" which contribute to generalized cellular functions and cellular survival, rather than to the determination of what

makes one cell-type different from.another. Moreover, highly expressed genes have generally already been isolated and characterized.

Unlike the housekeeping genes, non-methylated but normally-methylated, DNA fragments are present in just a small proportion of cells. Therefore, it is very difficult to

5   detect or isolate those fragments and to distinguish them from non-methylated housekeeping genes. The present invention solves this problem by using a combination of DNA amplification and hybridization/subtraction techniques. The amplification steps enrich the genomic DNA sample with small DNA fragments that have both 5' and 3' ends cut by the methylation-sensitive enzyme. The hybridization/subtraction steps eliminate fragments

10   encoding non-methylated housekeeping genes and fragments which are cut on only one end by the methylation-sensitive enzyme. Several rounds of amplification and hybridization/subtraction can be used if a discrete number of fragments are not isolated with the first round.

The fragments isolated by MDD are cut on both ends by the methylation-sensitive

15   master enzyme because only fragments cut on both ends with the master enzyme are efficiently amplified. This is accomplished by using primers during DNA amplification that recognize only the master-enzyme-cut DNA ends. Such primer specificity is created by adding adaptors to the DNA ends of a sample cut with both the master enzyme and a partner restriction enzyme. The adaptors have unique sequences which hybridize with the

20   DNA amplification primers. The partner enzyme is used to generate a mixture of DNA fragments of an appropriate size and complexity for DNA amplification but also creates a population of DNA molecules which will not accept the adaptors because they have partner-enzyme-cut ends.

Similarly, for detecting and isolating fragments which contain a mutation, DNA

25   fragments are cut with a master enzyme and a partner (detector) enzyme, where the partner (detector) enzyme may cut normal DNA but which will not cut mutated DNA. The master enzyme normally used for MDD is a methylation-sensitive enzyme which recognizes and cuts CG-rich sites in genomic DNA. However, for detecting and isolating fragments that contain a mutation, nonmethylation sensitive master enzymes can also be used. Any

30   restriction enzyme can be used as a master restriction enzyme for detecting and isolating

mutations so long as it cuts at a different sequence than the detector enzyme. However, a master enzyme which cuts in CG-rich regions is preferred because it selects for regions associated with genes while the detector enzyme identifies and permits isolation of the mutation by cutting only the nonmutant DNA fragments

To confirm whether fragments isolated by the present MDD methods can detect differentially methylated or mutated DNA sites, they are tested as probes on genomic DNAs which are cut with methylation-sensitive or mutant-detector enzymes, respectively. If the probe detects a different size fragment in DNA which is known, or suspected to be, hypomethylated, hypermethylated or mutated than in normal DNA, then the probe can detect differentially methylated or mutated DNA sites.

DNA fragments isolated by the present methods can therefore constitute probes, for example, for detecting hypomethylation, hypermethylation or mutation in DNA fragments which are normally methylated.

In particular, the present invention provides a method of isolating probes to detect hypomethylation in a CNG triplet of DNA which includes:

a) cleaving a tester sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

b) amplifying the cleaved tester sample using primers that recognize DNA ends cut by the master restriction enzyme, to generate a first tester amplicon;

c) melting and hybridizing the first test amplicon with a driver DNA to generate a tester-tester product and a tester-driver product;

d) adding nucleotides to the ends of the tester-tester product and the tester-driver product to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

e) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second tester amplicon;

f) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect hypomethylation a CNG triplet of DNA;

wherein the master restriction enzyme cleaves a nonmethylated CNG DNA sequence but does not cleave a methylated CNG DNA sequence;

wherein the partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein the driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by the master restriction enzyme.

In another embodiment, the master-enzyme-cut DNA ends are distinguished from the partner-enzyme-cut DNA ends by ligating a first set of adaptors onto master-enzyme-cut DNA ends of the cleaved tester sample to generate a first-tester amplification template. This first-tester amplification template is then amplified to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors. Hence, only DNA fragments which were cut by the master restriction enzyme will be amplified. The first adaptors are then cleaved off the first-tester amplicon and a second set of adaptors is ligated onto master-enzyme-cut DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends. As before, this tester DNA is melted and hybridized with a large molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product. The DNA ends of these products are made blunt by adding nucleotides and the blunt-ended tester-driver product is amplified the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends. This generates a second-tester amplicon from which a discrete DNA fragments can be isolated as probes to detect the hypomethylation in a CNG triplet of DNA.

As used herein, hypomethylation means that at least one cytosine in a CG or CNG di- or tri-nucleotide site in genomic DNA of a given cell-type does not contain CH_3 at the fifth position of the cytosine base. Cell types which may have hypomethylated CGs or CCGs include any cell type which may be expressing a non-housekeeping function. This includes both normal cells that express tissue-specific or cell-type specific genetic functions, as well as tumorous, cancerous, and similar cell types. Cancerous cell types and conditions

which can be analyzed, diagnosed or used to obtain probes by the present methods include Wilm's cancer, breast cancer, ovarian cancer, colon cancer, kidney cell cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, and hepatoblastoma.

Genomic DNA samples can be obtained from any mammalian body fluid, secretion, cell-type, or tissue, as well as any cultured cell or tissue.

The present MDD technique involves preparation of a driver DNA and a tester DNA. A driver DNA is genomic DNA from a "normal" cell type. A tester DNA is genomic DNA from a "test" cell type which may have mutated DNA sites, hypermethylated DNA sites or non-methylated DNA sites in place of the normal DNA at that site. Such normal cell types are cells of the same tissue-type as the "test" cell type, but which are normally methylated and are not mutated. Test cell types include cells that express tissue or cell-specific functions, cancer cells, tumor cells and the like. For example, if breast cancer cells are selected as the test cell type, the normal cell type would be breast cells that are not diseased or cancerous. Preferably, test and normal cell types are isolated from the same person.

After isolation, the tester DNA and the driver DNA are separately cut with both the master and partner restriction enzymes. As provided by the present invention, the master restriction enzyme cleaves a CNG DNA sequence. The master restriction enzyme produces sticky ends, rather than blunt ends, when it cuts. In general, the master enzyme used in the present methods cleaves only non-methylated CNG DNA sequences and does not cleave methylated CNG DNA sequences. However, for selected applications a non-methylation sensitive master enzyme can be used, for example, for detecting and isolating mutations. For detecting and isolating hypomethylated DNA sites, a methylation-sensitive master enzyme should be used, for example, the *Msp* I restriction enzyme which recognizes and cleaves DNA at nonmethylated CCGG but will not cleave the CCGG sequence when the outer cytosine is methylated. Master restriction enzymes contemplated by the present invention include *Msp* I, *Bsi*S I, *Hin*2 I and the like. In a preferred embodiment, *Msp* I is used as the master restriction enzyme for detecting and isolating a DNA fragment containing a hypomethylated site.

According to the present invention, a partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme. As used herein, the complexity of mixture of DNA fragments is the number of different DNA fragments within a selected size range. When mammalian DNA is used, the present methods can effectively screen a DNA sample containing about 1% to about 25% of the mammalian genome. In a preferred embodiment, a DNA sample has about 5% to about 15% of the total DNA sequences in the mammalian genome. In an especially preferred embodiment, the DNA sample has about 10% of the total DNA sequences in the mammalian genome.

According to the present invention, the size range of DNA fragments in a sample of genomic DNA is that size range which can be amplified by in vitro DNA amplification. When polymerase chain reaction ("PCR") is used, that size range is about 2000 base pairs (bp) to about 75 bp. Preferably the selected size range is about 1500 bp to about 100 bp. More preferably the selected size is about 1000 bp to about 150 bp.

As used herein, the preferred size is generated by the action of both the master and partner restriction enzymes. Hence, if the partner restriction enzyme generates a population of genomic DNA fragments which are about 1000 bp or smaller, but, in combination with the master restriction enzyme a population of fragments which are less than 50 bp is generated, a different set of partner and master enzymes is preferably used.

Also according to the present invention, the partner restriction enzyme cleaves DNA to produce ends that are neither homologous nor complementary to a sticky end produced by the master restriction enzyme.

Any convenient restriction enzyme is used as a partner restriction enzyme to identify CNG sites that are hypomethylated, so long as the appropriate complexity and size of DNA fragments are generated. For convenience, a partner restriction enzyme that operates under the same conditions as the master restriction enzyme may be used. For example, *Mse* I is a good partner restriction enzyme which can be used under the same conditions as the preferred *Msp* I master restriction enzyme. When the preferred *Msp* I master restriction enzyme is used, the partner restriction enzyme preferably recognizes and cuts a 4-nucleotide

sequence. Moreover, in a preferred embodiment, the partner restriction enzyme cuts an AT-rich site or generates a blunt end to avoid any ligation of adaptors, which are designed to attach to master enzyme cut ends, onto the ends generated by the partner enzyme. Preferred partner restriction enzymes include *Mse* I, *Sau*3 A, *Rsa* I, *Tsp*E I, *Mae* I, *Nia* III,

5   *Dpn* I and the like.

Cleavage methods and procedures for selecting restriction enzymes for cutting DNA at specific sites are known to the skilled artisan. For example, many suppliers of restriction enzymes provide information on conditions and types of DNA sequences cut by specific restriction enzymes, including New England BioLabs, ProMega Biochem,

10   Boehringer-Mannheim and the like. Sambrook et al. (1989) provide a general description of methods for using restriction enzymes and other enzymes.

After cleaving the DNA with the master and partner restriction enzymes, a set of adaptors is hybridized and ligated onto the sticky DNA ends produced by the master restriction enzyme. The adapters are selected so that they will ligate to the CG-rich ends of

15   the DNA cut by the master restriction enzyme but not to the ends of the fragments that were cut with the partner restriction enzyme. Because only the DNA ends of fragments in the tester and driver genomic DNA samples will be single-stranded, adaptors need not be longer than about 12 nucleotides to specifically recognize and hybridize to the master enzyme-cut ends. However, the adaptors are chosen not only to ligate to DNA ends cut by

20   the master restriction enzyme, but also to be a good size and DNA sequence to act as recognition sites for primers used during DNA amplification. Adaptors are chosen so that, when the mixture of master and partner enzyme cut DNA fragments is amplified using primers that hybridize to those adaptors, only those fragments that have the adaptor and thus were cut with the master restriction enzyme will be amplified. Hence, after

25   hybridization and ligation the adaptor ends should have a unique sequence and be of sufficient length to be a unique recognition site for primers selected for in vitro DNA amplification.

Preferred sets of adaptors of the present invention include those of SEQ ID NO: 1-6. As used herein:

the MSA24 adaptor has SEQ ID NO:1 --

CTCGTCGTCAGGTCAGTGCTTCAC

the MSA12 adaptor has SEQ ID NO:2 -- CGGTGAAGCACT

the MSB24 adaptor has SEQ ID NO:3 --

TAGAGCCACGTAGCTGCTGTAGTC

the MSB12 adaptor has SEQ ID NO:4 -- CGGACTACAGCA

the MSC24 adaptor has SEQ ID NO:5 --

ACCGTGGACTGGATAGGTTCAGAC

and the MSC12 adaptor has SEQ ID NO:6 -- CGGTCTGAACCT.

In one embodiment, the SEQ ID NOS: 1 and 2 (MSA24 and MSA12) adaptors are single-stranded and are used as the first set of adaptors. These single-stranded adaptors are annealed and hybridized onto DNA ends cut with the master restriction enzyme and then ligated to those DNA ends by incubation with ligase. Both the tester and driver DNAs receive the first set of adaptors. Thus, the master-cut ends of both the tester and driver DNAs will have adaptors while DNA ends cut with the partner enzyme, from either the tester or driver, will have no such adaptors.

The tester and driver samples are then separately amplified by known DNA amplification procedures using primers that hybridize to the adaptors. For example, when the SEQ ID NO:1 and 2 adaptors are used, an oligonucleotide having SEQ ID NO:1 can be used as a primer for DNA amplification of only the master enzyme cut DNA fragments. Such DNA amplification generates a tester amplicon and a driver amplicon. In both the tester and driver amplicons, only fragments which are cut with the master restriction enzyme will be amplified.

DNA amplification procedures for use in the present MDD methods include any in vitro amplification procedure which provides sequence-specific synthesis of a nucleic acid fragment relative to the complex bulk of nucleic acid present in a sample. The specificity of the process is determined by enzymatic recognition of a specific sequence or by the oligonucleotide primers which are capable of hybridizing only with selected adaptors that are added onto the ends of master restriction cut DNA fragments.

In vitro DNA amplification.techniques are known in the art. A review of such

techniques can be found in Kwoh and Kwoh (1990). In vitro nucleic acid amplification

techniques include polymerase chain reaction (PCR), transcription-based amplification

system (TAS), self-sustained sequence replication system (3SR), ligation amplification

5   reaction (LAR), ligase-based amplification system (LAS), Qß RNA replication system and

run-off transcription.

PCR is a preferred method for DNA amplification. PCR synthesis of DNA

fragments occurs by repeated cycles of heat denaturation of DNA fragments, primer

annealing onto the adaptor ends of the master-cut DNA fragment, and primer extension.

10   These cycles can be performed manually or, preferably, automatically. Thermal cyclers

such as the Perkin-Elmer Cetus cycler are specifically designed for automating the PCR

process, and are preferred. The number of cycles per round of synthesis can be varied

from 2 to more than 50, and is readily determined by considering the source and amount of

the nucleic acid template, the desired yield and the procedure for detection of the

15   synthesized DNA fragment.

PCR techniques and many variations of PCR are known. Basic PCR techniques

are described by Saiki et al. (1988) and by U.S. Patent Nos. 4,683,195, 4,683,202 and

4,800,159, which are incorporated herein by reference.

The conditions generally required for PCR include temperature, salt, cation, pH and

20   related conditions needed for efficient copying of the master-cut fragment. PCR conditions

include repeated cycles of heat denaturation (i.e. heating to at least about 95°C) and

incubation at a temperature permitting primer: adaptor hybridization and copying of the

master-cut DNA fragment by the amplification enzyme. Heat stable amplification enzymes

like the pwo, Thermus aquaticus or Thermococcus litoralis DNA polymerases are

25   commercially available which eliminate the need to add enzyme after each denaturation

cycle. The salt, cation, pH and related factors needed for enzymatic amplification activity

are available from commercial manufacturers of amplification enzymes.

As provided herein an amplification enzyme is any enzyme which can be used for in

vitro nucleic acid amplification, e.g. by the above-described procedures. Such amplification

30   enzymes include pwo, Escherichia coli DNA polymerase I, Klenow fragment of E. coli

DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, Thermus aquaticus (Taq) DNA polymerase, Thermococcus litoralis DNA polymerase, SP6 RNA polymerase, T7 RNA polymerase, T3 RNA polymerase, T4 polynucleotide kinase, Avian Myeloblastosis Virus reverse transcriptase, Moloney Murine Leukemia Virus reverse transcriptase, T4

5   DNA ligase, E. coli DNA ligase or Qβ replicase. Preferred amplification enzymes are the pwo and Taq polymerases. The pwo enzyme is especially preferred because of its fidelity in replicating DNA.

The length of the primers for use in MDD depends on several factors including the nucleotide sequence and the temperature at which these nucleic acids are hybridized or

10   used during in vitro nucleic acid amplification. The considerations necessary to determine a preferred length for an oligonucleotide primer are well known to the skilled artisan. For example, as is known to the skilled artisan, the length of a short nucleic acid or oligonucleotide can relate to its hybridization specificity or selectivity. Because the tester and driver samples contain a complex mixtures of nucleic acids, primers which are shorter

15   than about 12 nucleotides may hybridize to more than one site in the test genomic DNA, and accordingly would not have sufficient hybridization selectivity for amplifying only the master-cut fragments. However, a 12- to 15-nucleotide sequence is generally represented only once in a mammalian genome (Sambrook, 1989, Vol. 2, pp. 11.7-11.8). Accordingly, to eliminate amplification of fragments which do not have adaptors, primers are chosen

20   which are generally at least about 14 nucleotides long.

Preferably, the present primers are at least 16 nucleotides in length. More preferred primers are at least 17 nucleotides in length (Sambrook, pp. 11.7- 11.8).

After amplification, the first set of adaptors are removed from the ends of the tester and driver amplicons by cleavage with the master restriction enzyme. Cleavage with the

25   master restriction enzyme preserves the original ends of the fragments and removes all DNA which was not present in the tester DNA as originally isolated.

A second set of adaptors are ligated to the tester amplicon, but not to the driver amplicon. This generates a second-adaptor-tester sample. The second set of adaptors is selected so that they do not have the same sequence as the first set of adaptors and so that

30   they hybridize, and hence ligate, only to DNA ends cut by the master restriction enzyme.

The second set of adaptors should also form a good recognition site for primers which are used during DNA amplification.

During DNA amplification, non-methylated DNA fragments which are normally present in all cell types, and which may not be of interest, will also be amplified because they will have adaptors on their ends. At least one round of subtraction/hybridization followed by DNA amplification is used by the present MDD methods to remove these nonmethylated DNA fragments which likely encode "housekeeping" functions.

In general, the present MDD techniques remove "housekeeping" sequences by hybridizing the amplified tester DNA with a large molar excess of amplified driver DNA and amplifying only the tester-tester hybrids which have perfectly hybridized ends because those perfectly hybridized ends are the only sites recognized by the primers used for DNA amplification.

In particular, the amplified tester DNA (called the first test amplicon or the second-adaptor-tester) is mixed with a large molar excess of amplified driver DNA. The mixture is denatured, then renatured. As used herein, a large molar excess of driver to tester DNA is about a 10-fold to about a 10,000-fold molar excess of driver DNA to tester DNA. In a preferred embodiment the tester DNA is melted and hybridized with about a 10-fold to about a 1000-fold molar excess of a driver DNA. In a more preferred embodiment, the tester DNA is melted and hybridized with about a 100-fold molar excess of a driver DNA. Stringent hybridization conditions are preferably used during renaturation to maximize the formation of fragments with few mismatches.

The skilled artisan can vary the hybridization conditions used during MDD to achieve optimal hybridization of unique tester DNA fragments to their tester homolog. Conditions for achieving optimal hybridization are known to the skilled artisan and generally include temperature and salt concentrations permitting selective hybridization between two highly homologous DNA fragments, e.g. stringent hybridization conditions. Stringent conditions permit little or no detectable hybridization between mismatched driver or tester fragments, that is between fragments that have dissimilar sequences, particularly at the ends. Hybridization techniques are described, for example, in Sambrook et al. (1989).

The ends of the renatured fragments are then filled in using a DNA polymerase, e.g. Taq or pwo polymerase. Such a "filling in" reaction makes the DNA fragments blunt-ended but does not eliminate any internally mismatched or unhybridized regions. To eliminate such mismatched hybrids, the renatured fragments may be treated with a nuclease that will

5   remove single-stranded regions and mismatched nucleotides in the renatured fragments. For example, S1 nuclease, mung bean nuclease and the like can be used to eliminate mismatched hybrids. The nuclease will also cleave off mismatched nucleotides at the ends of either the tester or the driver fragment. Hence, mismatched hybrids will have ends that do not contain second adaptor ends. Only the single-stranded tester fragments with second

10   adaptor ends that faithfully hybridize to an exact tester homolog will end up with a second adaptor end.

The subtraction/hybridization mixture is then amplified by in vitro DNA amplification procedures using primers that hybridize to the second adaptor ends. Hence, only the tester DNA fragments with second adaptor ends are amplified. Any tester DNA which

15   hybridized with driver DNA will not be amplified. A large excess of driver DNA is used to promote formation of hybrids that are commonly found in both the tester and driver samples. The result is a second tester amplicon with fragments which were uniquely unmethylated in the original tester DNA sample.

DNA fragments cut on only one end with the master restriction enzyme ("single

20   master-cut fragments") are eliminated by the present invention is several ways. First, only one end of those fragments will ligate to an adaptor. This means that, during any amplification step, the single master-cut fragments will not be amplified as quickly as the "double master-cut fragments" which have adaptors at both ends. As is known to the skilled artisan, fragments which hybridize with amplification primers at both ends are

25   amplified geometrically, while DNA fragments which hybridize with amplification primers at only one end are amplified arithmetically. Second, during the subtraction/ hybridization step those single master-cut fragments may also hybridize to their driver homologs because the driver is present in vast molar excess. This imperfect hybrid can be eliminated by treatment with nuclease, so it will not be amplified during the subsequent DNA amplification step.

Several rounds of amplification and subtraction/hybridization can be used to isolate discrete DNA fragments which can be used as probes to detect hypomethylation, hypermethylation and mutation in specific regions of the genome. Discrete DNA fragments isolated by the present MDD methods which can be used to detect hypomethylation,

5   hypermethylation and mutations in genomic DNA samples from different cell types include DNA fragments having SEQ ID NO: 7-10.

While discrete DNA fragments which detect hypomethylation also frequently detect hypermethylation, the MDD method is readily adapted to permit isolation of fragments which will specifically detect hypermethylation in identified sites of the genome. This is

10   accomplished merely by using normal cell DNA as the source of DNA for the hypermethylation probe. In this case, the normally unmethylated DNA will be cut by the methylation-sensitive master enzyme while the hypermethylated test DNA will not be cut. To eliminate DNA fragments which are common to the normal and test DNA samples, the normal cell DNA is hybridized with an excess of test cell DNA (now called "test driver

15   DNA"). Hence, isolation of hypermethylation probes generally involves the same manipulations, as used for isolation of hypomethylation probes, but the roles of the normal cell DNA and the test cell DNA are reversed.

Also for detecting and isolating DNA fragments that have a hypermethylated CNG, a methylation-sensitive enzyme can be used as a partner enzyme and a non-methylation

20   sensitive enzyme as a master enzyme. When this modification is used, the adaptors are still designed to attach to the ends of DNA fragments cut by the non-methylation sensitive master restriction enzyme.

In particular, the present invention provides a method of isolating a probe to detect hypermethylation in a CNG triplet of DNA which includes:

25   a) cleaving a normal sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved normal sample;

b) amplifying the cleaved normal sample using primers that recognize DNA ends cut by the master restriction enzyme, to generate a first normal amplicon;

c) melting and hybridizing the first normal amplicon with a test driver DNA to

30   generate a normal-normal product and a normal-driver product;

d) adding nucleotides to the ends of the normal-normal product and the normal-driver product to make a blunt-ended normal-normal product and a blunt-ended normal-driver product;

e) amplifying the blunt-ended normal-normal product and the blunt-ended normal-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second normal amplicon;

f) isolating a discrete DNA fragment from the second normal amplicon as a probe to detect the hypermethylation in a CNG triplet of DNA;

wherein the test driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by the master restriction enzyme.

The normal sample of genomic DNA is obtained by isolating genomic DNA from a "normal" cell type, and the test driver DNA is obtained by isolating genomic DNA from a "test" cell type which may have hypermethylated DNA sites in place of the normal non-methylated DNA at that site. Here, the test cell type is a cancerous or tumorous cell type.

In a further embodiment, the present MDD techniques can be used to identify and isolate mutations in genomic DNA and it has particular utility for identifying mutations which are near the promoters or coding regions of genes. Nonmethylation-sensitive master restriction enzymes can be used for isolating mutations. However, the preferred master enzymes recognize CNG DNA sequences including CpG dinucleotides which are frequently methylated in human genome. Condensed CpG islands appear in gene promoter regions, or even in the first exons. Because CpG islands appear in gene promoter regions, and in the first few exons of many genes, MDD-identified lesions will often be close to genes which cause disease.

According to the present invention, MDD can be used to isolate probes for detecting any type of genomic lesion. For example, MDD can be used to identify fragments which have point mutations, deletions, insertions, amplifications, rearrangements, and other mutations.

MDD is readily adapted for isolating mutations in DNA sequence by using a partner enzyme, now called a detector enzyme, that cuts normal DNA but which will not cut the

mutated DNA. For example, a point mutation in tester DNA can be identified when a cleavage site recognized by the detector enzyme is eliminated. When cut with master and detector restriction enzymes, a fragment with both ends cleaved by the master enzyme will be present in the tester DNA. However, the detector enzyme cleaves this fragment in the

5   driver DNA, because normal DNA has no point mutation. As a result, the fragment containing the point mutation will be isolated after performing the present MDD amplification and subtraction/hybridization steps. The point mutation can be easily identified by comparing the DNA sequence of the isolated mutant fragment with the DNA sequence of the homolog isolated from a matched normal genomic DNA. Moreover, the mutant

10   tester fragment will contain the entire mutation -- rather than just part of the mutation as would occur if a detector enzyme were used that cut the mutation rather than the normal DNA site.

Similarly, MDD is readily adapted for isolating deletion mutations in DNA simply by using tester DNA as the driver DNA and normal DNA as the source of the intact fragment

15   whose homolog in the tester DNA has a deletion. The deletion mutant fragment can be isolated by virtue of its partial homology with the normal DNA fragment.

In particular, the present invention provides a method of identifying a probe to detect a mutation in a tester sample of genomic DNA which includes:

a) cleaving a tester sample of genomic DNA with both a master restriction enzyme

20   and a detector restriction enzyme to generate a cleaved tester sample;

b) ligating a first set of adaptors onto master enzyme cut DNA ends of the cleaved tester sample to generate a first-tester amplification template;

c) amplifying the first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors;

25   d) cleaving off the first adaptors from the first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) melting and hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product

30   and tester-driver product;

f) adding nucleotides onto DNA ends present in the mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

h) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect a mutation in the tester sample of genomic DNA;

wherein the detector restriction enzyme cleaves a normal DNA site but a mutant DNA site to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein the driver DNA is cut with both the master restriction enzyme and the detector restriction enzyme and amplified using primers that recognize DNA ends cut by the master restriction enzyme.

Probes isolated by the present MDD techniques have at least about 14 nucleotides to about 2000 nucleotides.

It is well known in the art to adjust hybridization and wash solution contents and temperatures such that stringent hybridization conditions are obtained. Preferred stringent conditions are those that allow a probe to hybridize to sequences which are more than 90% complementary to the probe and not to those which are less than 70% complementary. Stringency depends on such parameters as the size and nucleotide content of the probe being utilized. See Sambrook et al. and other sources for general descriptions and examples.

For identifying mutations and isolating DNA fragments containing those mutations, selection of a detector restriction enzyme involves identifying an enzyme which will not cut mutated genomic DNA but which will cut normal DNA. The skilled artisan can readily perform this analysis, for example, by testing a plurality of restriction enzymes in the present MDD methods.

According to the present invention, a mutation, hypomethylation and hypermethylation at a specific site in genomic DNA can be detected by observing whether the size or intensity of a DNA fragment cut with a master or detector restriction enzyme is

the same in control and test samples. This can be done by cutting genomic DNA isolated from control and test tissue samples with the master or detector restriction enzyme, hybridizing a probe to the control and test DNAs and observing whether the two hybridization complexes are the same or different sizes or intensities.

In particular, the present invention provides a method of detecting whether a CNG triplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cells which includes:

a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;

b) cleaving the control-cell DNA and the test-cell DNA with a master restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;

c) preparing a probe from a DNA isolated by the methods of the present invention, for example, a DNA selected from the group consisting of SEQ ID NO:7-10;

d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and

e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex.

Similarly, the present invention provides a method of detecting whether a DNA site is mutated in a genomic DNA present in a test sample of cells which includes:

a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;

b) cleaving the control-cell DNA and the test-cell DNA with a detector restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;

c) preparing a probe from a discrete DNA isolated by the methods of the present invention;

d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and

e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex;

wherein the detector restriction enzyme does not cleave a mutated DNA site but does cleave a corresponding nonmutated DNA site.

Hybridization techniques for detecting specific genomic DNA fragments are known in the art and include solid-phase-based hybridization and solution hybridization which use any of the known reporter molecules. Detailed methodology for gel electrophoretic and nucleic acid hybridization techniques can be found in Sambrook et al.

Preferred nucleic acid probes for detecting hypomethylation and hypermethylation have SEQ ID NOS:7-10.

According to the present invention, a probe isolated by the subject methods, can be labeled by any procedure known in the art, for example by incorporation of nucleotides linked to a "reporter molecule".

A "reporter molecule", as used herein, is a molecule which provides an analytically identifiable signal allowing detection of a hybridized probe. Detection may be either qualitative or quantitative. Commonly used reporter molecules include fluorophores, enzymes, biotin, chemiluminescent molecules, bioluminescent molecules, digoxigenin, avidin, streptavidin or radioisotopes. Commonly used enzymes include horseradish peroxidase, alkaline phosphatase, glucose oxidase and β-galactosidase, among others. Enzymes can be conjugated to avidin or streptavidin for use with a biotinylated probe. Similarly, probes can be conjugated to avidin or streptavidin for use with a biotinylated enzyme. The substrates to be used with these enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. For example, p-nitrophenyl phosphate is suitable for use with alkaline phosphatase reporter molecules; for horseradish peroxidase, 1,2-phenylenediamine, 5-aminosalicylic acid or tolidine are commonly used.

Incorporation of a reporter molecule into a DNA probe can be by any method known to the skilled artisan, for example by nick translation, primer extension, random oligo priming, by 3' or 5' end labeling or by other means (see Sambrook, 1989).

In another embodiment, the present invention provides one or more compartmentalized kits for detection of hypomethylation. A first kit has a receptacle containing at least one of the present isolated probes. Such a probe may be a nucleic acid fragment which is methylated at a CNG trinucleotide in the genomic DNA of normal cells

but which is not methylated in the genomic DNA of cancerous cells. Such a probe may be specific for a DNA site that is normally methylated but which is nonmethylated in certain cell types. Similarly, such a probe may be specific for a DNA site that is abnormally methylated or hypermethylated in certain cell types. Finally, such a probe may identify a specific DNA mutation. By specific for a DNA site is meant that the probe is capable of hybridizing to the DNA sequence which is mutated or abnormally methylated, or is capable of hybridizing to DNA sequences adjacent to said mutated or abnormally methylated DNA sequences. The probe provided in the present kits may have a covalently attached reporter molecule.

A second kit has at least one receptacle containing a reagent for MDD. Such a reagent can be, for example, a preferred master restriction enzyme, a preferred partner restriction enzyme, a preferred adaptor which can be ligated to DNA cut with the preferred master restriction enzyme, a primer that hybridizes to a preferred adaptor, and the like.

Using the MDD methods of the present invention, several methyl-polymorphic markers or sites in the human genome have been identified and DNA fragments encoding these marker sites have been isolated from different cell types. In particular, the present MDD methods have been used to identify hypomethylated sites which act as markers indicating whether a patient has cancer. These sites were identified using B-cells from patients with Chronic Lymphocytic Leukemia (CLL) and using tumor cells from breast cancer patients.

Moreover, for the first time the present methods and probes experimentally prove that methylated external cytosine at CpCpG triplet sequences do exist in the human genome, and that alteration of their methylation patterns is related to human cancer development. According to the present invention, methylation of the first base of CpCpG triplets is not a random event. Instead, it is a site- and tissue-specific event.

The alteration of the methylation pattern in CpCpG triplets may be a key, and common event, in the development of neoplasia. Imbalanced methylation at CpCpG triplet may play at least two roles in tumorigenesis:

1) DNA hypomethylation may cause some proto-oncogene expression or DNA hypermethylation may silence a tumor supressor which contributes to neoplastic growth; and

2) DNA hypomethylation may change chromatin structure, and induce abnormalities in chromosome pairing and disjunction. Such structural abnormalities may result in genomic lesions, such as chromosome deletions, amplifications, inversions, mutations, and translocations, all of which are found in human genetic diseases and cancer.

Hence, the present MDD methods have broad utility for identifying differentially methylated sites at CpCpG triplet sequences in the human genome; for mapping hypo- and hyper-methylation sites which are related to disease development; for understanding the role(s) of DNA methylation in normal cell genomic DNA imprinting, differentiation, and development; for understanding tumorigenesis; for diagnosing and monitoring the prognosis of disease; and for searching for proto-oncogene(s) and recessive oncogene(s). MDD, as a comprehensive molecular genetic technique, enables the accumulation of knowledge necessary to help better understand the etiology of human genetic diseases and cancer.

The methods of the present invention permit identification and isolation of genomic nucleic acids which are differentially methylated or differentially expressed among tissues of different types. Using the nucleotide sequences provided, oligonucleotides can be constructed which are useful as reagents for detecting the susceptibility of particular DNA sequences in a test sample to cleavage by restriction endonucleases. Such susceptibility to cleavage is affected by factors which include methylation of particular bases and the substitution of bases through mutation. For example, from the sequences provided, it is possible to generate nucleic acid probes to determine the sizes and relative amounts of particular DNA species which result when DNA recovered from a test sample is digested with selected sequence specific endonucleases.

The genomic nucleic acids isolated by the present invention, or oligonucleotides obtained therefrom, can also be used to detect and isolate expression products, for example, mRNA, complementary DNAs (cDNAs) and proteins derived from or homologous to the present genomic nucleic acids. In one embodiment, the expression of mRNAs homologous to the nucleic acids of the present invention is detected, for example, by Northern analysis, reverse transcription or amplification by PCR. Such procedures permit detection of mRNAs in a variety of tissue types or at different stages of development.

The subject nucleic acids which are expressed in a tissue-specific or a developmental-stage-specific manner are useful as tissue-specific markers or for defining the developmental stage of a sample of cells or tissues. One of skill in the art can readily adapt the present nucleic acids and oligonucleotides for these purposes. Useful hybridization conditions for

5   probes and primers made according to the present invention are readily determinable by those of skill in the art. Preferred stringent hybridization conditions are those which allow hybridization of nucleic acids which are greater than 90% homologous but which prevent hybridization of nucleic acids which are less than 70% homologous.

In another embodiment of the present invention, cDNAs homologous to the present

10   nucleic acids can be isolated by procedures readily available to those of skill in the art. For example, cDNA libraries can be obtained from the tissues or cells which express the present genomic nucleic acids. Those cDNA libraries can then be probed with a genomic nucleic acid isolated by the methods of the present invention, or an oligonucleotide fragment thereof. Again, one of skill in the art can readily isolate such cDNAs using the genomic

15   nucleic acids of the present invention.

The present genomic and cDNA nucleic acids can be used to prepare oligonucleotides. In one embodiment, oligonucleotides having twelve nucleotides or more can be prepared which hybridize specifically to the present genomic nucleic acids and cDNAs and allow detection and isolation of unique nucleic sequences by hybridization.

20   Sequences of at least 15-20 nucleotides are preferred and are selected from regions which lack homology to other known sequences. Sequences of 20 or more nucleotides which lack such homology are optimal.

Genomic nucleic acids isolated by the present procedures include *BR50* (SEQ ID NO:10), *BR104* (SEQ ID NO:7) and *BR254* (SEQ ID NO:8), clones which are

25   hypomethylated in breast cancer cells. A cDNA clone identified as *TSP50* (SEQ ID NO:14) was obtained by virtue of its homology to the *BR50* (SEQ ID NO:10) genomic clone. Each of these nucleic acids is useful for detecting hypomethylation and expression of homologous nucleic acids in a variety of normal and cancerous cells or tissues. Each of these nucleic acids is also useful for identifying homologous genomic and cDNA nucleic

30   acids.

Oligonucleotides useful as probes for screening of samples by hybridization assays or as primers for amplification can be packaged into kits. Such kits usually contain the probes or primers in a premeasured or predetermined amount, as well as other suitably packaged reagents and materials needed for the particular hybridization or amplification protocol.

In another embodiment, the present genomic and cDNA nucleic acids can be placed in expression vectors capable of expressing an encoded protein, polypeptide or peptide. Nucleic acids encoding a desired protein or polypeptide are derived from cDNA or genomic clones using conventional restriction digestion or by synthetic methods, and are ligated into vectors from which they may be expressed. Vectors may, if desired, contain nucleic acids encoding portions of other proteins, thereby providing a fusion protein. For example, portions of the β-galactosidase or superoxide dismutase coding region may be placed in frame with the present coding regions to provide an expression vector which encodes a fusion protein.

Expression vectors include plasmids designed for the expression of proteins or polypeptides fused to or within bacterial phage coat proteins. The DNA encoding the desired protein or polypeptide, whether in a fusion, premature or mature form, may be ligated into expression vectors suitable for any host. The DNA encoding the desired polypeptide may also contain a signal sequence to permit secretion from the intended host. Both prokaryotic and eukaryotic host systems are contemplated.

The present proteins, polypeptides and peptides, whether in a premature, mature or fused form, are isolated from lysed cells, or from the culture medium, and are purified to the extent needed for the intended use. One of skill in the art can readily purify these proteins, polypeptides and peptides by any available procedure. For example, purification may be accomplished by salt fractionation, size exclusion chromatography, ion exchange chromatography, reverse phase chromatography, affinity chromatography and the like.

The present invention further contemplates antibodies which can bind to the present proteins, polypeptides and peptides. Such antibodies preferably bind to unique antigenic regions or epitopes in the present proteins, polypeptides and peptides. Unique antigenic regions or epitopes of a protein are generally small, often eight to ten amino acids or less in

length. An epitope or antigenic region may be defined by a sequence of as few as 5 amino acids.

Epitopes and antigenic regions useful for generating antibodies can be found within the present proteins, polypeptides or peptides by procedures available to one of skill in the art. For example, short, unique peptide sequences can be identified in the present proteins and polypeptides that have little or no homology to known amino acid sequences. Preferably the region of a protein selected to act as a peptide epitope or antigen is not entirely hydrophobic; hydrophilic regions are preferred because those regions likely constitute surface epitopes rather than internal regions of the present proteins and polypeptides. These surface epitopes are more readily detected in samples tested for the presence of the present proteins and polypeptides.

Peptides can be made by any procedure known to one of skill in the art, for example, by using in vitro translation or chemical synthesis procedures.

Short peptides which provide an antigenic epitope but which by themselves are too small to induce an immune response may be conjugated to a suitable carrier. Suitable carriers and methods of linkage are well known in the art. Suitable carriers are typically large macromolecules such as proteins, polysaccharides and polymeric amino acids. Examples include serum albumins, keyhole limpet hemocyanin, ovalbumin, polylysine and the like. One of skill in the art can use available procedures and coupling reagents to link the desired peptide epitope to such a carrier. For example, coupling reagents can be used to form disulfide linkages or thioether linkages from the carrier to the peptide of interest. If the peptide lacks a disulfide group, one may be provided by the addition of a cysteine residue. Alternatively, coupling may be accomplished by activation of carboxyl groups.

The minimum size of peptides useful for obtaining antigen specific antibodies can vary widely. The minimum size must be sufficient to provide an antigenic epitope which is specific to the protein or polypeptide. The maximum size is not critical unless it is desired to obtain antibodies to one particular epitope. For example, a large polypeptide may comprise multiple epitopes, one epitope being particularly useful and a second epitope being immunodominant. Typically, antigenic peptides selected from the present proteins and polypeptides will range from 5 to about 100 amino acids in length. More typically,

however, such an antigenic peptide will be a maximum of about 50 amino acids in length, and preferably a maximum of about 30 amino acids. It is usually desirable to select a sequence of about 10, 12 or 15 amino acids, up to a maximum of about 20 or 25 amino acids.

Amino acid sequences comprising useful epitopes can be identified in a number of ways. For example, the entire protein sequence can be screened by preparing a series of peptides that together span the entire protein sequence. One of skill in the art can routinely test a few large polypeptides for the presence of an epitope showing a desired reactivity and also test progressively smaller and overlapping fragments to identify a preferred epitope with the desired specificity and reactivity.

Antigenic polypeptides and peptides are useful for the production of monoclonal and polyclonal antibodies. To obtain polyclonal antibodies, a selected animal is immunized with a protein or polypeptide. Serum from the animal is collected and treated according to known procedures. Polyclonal antibodies to the protein or polypeptide of interest may then be purified by affinity chromatography. Techniques for producing polyclonal antisera are will known in the art.

Mabs may be made by procedures available to one of skill in the art, for example, by fusing antibody producing cells with immortalized cells and thereby making a hybridoma. The general methodology for fusion of antibody producing B cells to an immortal cell line is well within the province of one skilled in the art.

Mabs of the invention can also be generated from mRNA extracted from bone marrow and spleen cells of immunized animals using combinatorial antibody library technology (Huse et al., 1989; Barbas et al., 1991; Clackson et al., 1991; Burton and Barbas, 1994). mRNA is extracted and reversed transcribed to cDNA. cDNAs encoding part or all of antibody heavy and light chains, particularly the heavy and light chain variable domains which together comprise the antigen binding site, are then amplified by PCR. An immunoglobulin (Ig) phage display library is then constructed, for example, by inserting the amplified cDNA encoding IgG heavy chain and the amplified cDNA encoding a light chain into a phage display vector (e.g., pComb3, Stratagene) such that one vector contains a cDNA insert encoding a heavy chain fragment in a first expression cassette of the vector,

and a cDNA insert encoding a light chain fragment in a second expression cassette of the vector. Following transformation of E. coli bacteria with the ligated vectors, the bacteria are infected with filamentous phage M13 helper phage using methods well known in the art. Phage particles thus produced display antibody heavy and light chain variable domains and contain the DNA sequences which encode those domains.

The phage particles are then used in a panning procedure which enriches for those having antibody binding sites which can bind to tsp50 protein. The enriched phage population can be subjected to further rounds of infection and panning to identify those with desired variable domain binding characteristics. The general methodologies involved in creating large combinatorial libraries using phage display technology is described and disclosed in U.S. Patent No. 5,223,409 issued Jun. 29, 1993 which is incorporated herein by reference to disclose and describe phage display methodology.

Antibodies of the invention include complete anti-tsp50 antibodies as well as antibody fragments and derivaties which comprise a tsp50 binding site. Antibody fragments include, but are not limited to Fab, F(ab')$_2$, Fv and scFv. Derivatives are macromolecules which comprise a tsp50 binding site linked to a functional domain. Functional domains include, but are not limited to signaling domains, toxins, enzymes and cytokines.

The present antibodies are useful for detecting the present proteins and polypeptides in specific tissues or in body fluids. Moreover, according to the present invention, detection of aberrantly expressed proteins or protein fragments is probative of a disease state. For example, expression of the present proteins or polypeptides from a hypomethylated genomic locus may indicate that the protein is being expressed in an inappropriate tissue or at an inappropriate developmental stage. Hence, the present antibodies are useful for detecting diseases associated with protein expression from improperly methylated genomic loci, for example, the types of cancers and genetic diseases contemplated herein.

Design of immunoassays is subject to a great deal of variation and a variety of these are known in the art. Immunoassays may use a monoclonal or polyclonal antibody reagent which is directed against one epitope of the antigen being assayed. Alternatively, a

combination of monoclonal or polyclonal antibodies may be used which are directed against more than one epitope. Protocols may be based, for example, upon competition, direct antigen-antibody reaction or sandwich type assays. Protocols may, for example, use solid supports or immunoprecipitation. The present antibodies can be labeled with a reporter

5   molecule for easy detection. Assays which amplify the signal from a bound reagent are also known. Examples include immunoassays which utilize avidin and biotin, or which utilize enzyme-labeled antibody or antigen conjugates, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents include antibodies directed against the present protein epitopes or antigenic regions,

10   packaged appropriately with the remaining reagents and materials required for the conduct of the assay, as well as a suitable set of assay instructions.

The major drawback of rodent Mabs is that although they may be administered to a patient for diagnostic or therapeutic purposes, they are recognized as foreign antigens by the immune system and are unsuitable for continued use. Antibodies which are not recognized

15   as foreign antigens by the human immune system have greater potential for both diagnosis and treatment. Methods for generating human and humanized antibodies are now well known in the art. For a review of antibody humanization and of the nomenclature applied to antibody domains and regions, see, e.g., Vaughan et al. (1998).

Chimeric antibodies can be constructed in which regions of a non-human Mab are

20   replaced by their human counterparts. A preferred chimeric antibody is one which has amino acid sequences which comprise one or more complementarity determining regions (CDRs) of a non-human Mab which binds to tsp50 grafted to human framework (FW) regions. Methods for producing such antibodies are well known in the art (See, e.g., Jones et al., 1996; Riechman et al., 1988, U.S. Patent 5,530,101 to Queen et al.). Amino acid

25   residues corresponding to CDRs and FWs are know to one of average skill in the art. For example, amino acid sequence alignments indicating the locations of CDRs and FWs can be found in Kabat et al. (1991). A variety of methods have been developed to preserve or to enhance affinity for antigen of antibodies comprising grafted CDRs. One way is to include in the chimeric antibody the foreign framework residues which influence the conformation of

30   the CDR regions. A second way is to graft the foreign CDRs onto human variable domains

EV 332 525 056 US   -41-

with the closest homology to the foreign variable region (Queen et al., 1989). Thus, grafting of one or more non-human CDRs onto a human antibody may also involve the substitution of amino acid residues which are adjacent to a particular CDR sequence or which are not contiguous with the CDR sequence but which are packed against the CDR in the overall

5   antibody variable domain structure and which affect the conformation of the CDR. Humanized antibodies of the invention therefore include human antibodies which comprise one or more non-human CDRs as well as such antibodies in which additional substitutions or replacements have been made to preserve or enhance binding characteristics.

Chimeric antibodies of the invention also include antibodies which have been

10   humanized by replacing surface-exposed residues to make the Mab appear human (Padlan, 1991). Antibodies to CD56 and CD19 have been humanized by this process with no loss of affinity (Roguska et al., 1994). Because the internal packing of amino acid residues in the vicinity of the antigen binding site remains unchanged, affinity is preserved. Substitution of surface-exposed residues of an anti-tsp50 antibody according to the invention for the

15   purpose of humanization does not mean substitution of CDR residues or adjacent residues which influence affinity for tsp50.

Chimeric antibodies can also include antibodies where some or all non-human constant domains have been replace with human counterparts. This approach has the advantage that the antigen binding site remains unaffected. However, significant amounts of

20   non-human sequences are present where variable domains are derived entirely from non-human antibodies. (see, e.g., LoBuglio et al., 1989; Elliot et al., 1994).

Antibodies of the invention include human antibodies which are antibodies which consist essentially of human sequences. Human antibodies can be obtained from phage display libraries wherein combinations of human heavy and light chain variable domains are

25   displayed on the surface of filamentous phage (See, e.g., McCafferty et al., 1990; Aujame et al., 1997). Combinations of variable domains are typically displayed on filamentous phage in the form of Fab's or scFvs. The library is screened for phage bearing combinations of variable domains having desired antigen binding characteristics. Preferred variable domain combinations are characterized by high affinity for tsp50. Preferred

30   variable domain combinations can also characterized by high specificity for tsp50 and little

cross-reactivity to other related antigens. By screening from very large repertoires of antibody fragments, $(2-10 \times 10^{10}$; see e.g., Griffiths et al., 1994) a good diversity of high affinity Mabs can be isolated, with many expected to have sub-nanomolar affinities for tsp50.

Alternatively, human antibodies can be obtained from transgenic animals into which unrearranged human Ig gene segments have been introduced and in which the endogenous mouse Ig genes have been inactivated (reviewed in Brüggemann and Taussig, 1997). Preferred transgenic animals contain very large contiguous Ig gene fragments that are over 1 Mb in size (Mendez et al., 1997) but human tsp50-specific Mabs of moderate affinity can be raised from transgenic animals containing smaller gene loci (See, e.g., Wagner et al., 1994; Green et al., 1994).

Transgenic animals capable of expressing only human Ig genes can also be used to raise polyclonal antiserum comprising antibodies solely of human origin.

Antibodies of the invention include those for which binding characteristics have been improved by direct mutation or by methods of affinity maturation. Affinity and specificity may be modified or improved by mutating CDRs and screening for antigen binding sites having the desired characteristics (See, e.g. Yang et al., 1995). CDRs can be mutated in a variety of ways. One way is to randomize individual residues or combinations of residues so that in a population of otherwise identical antigen binding sites, all twenty amino acids can be found at particular positions. Alternatively, mutations can be induced over a range of CDR residues by error prone PCR methods (See, e.g., Hawkins et al., 1992). Phage display vectors containing heavy and light chain variable region gene can be propagated in mutator strains of E. coli (See, e.g., Low et al., 1996). These methods of mutagenesis are illustrative of the many methods known to one of skill in the art.

Antibodies of the invention can be linked to agents which are useful for diagnosis or treatment. For diagnostic purposes, detectable agents are linked to an antibody which is specific for the antigen to be detected. The detectable label attached to the antibody for diagnosis may be an imaging agent for tumor imaging such as a short-lived radioisotope, for example ^{111}In, ^{125}I or [99m]Tc. A review of cancer imaging using antibodies against carcinoembryonic antigen is provided by Goldenberg (1992).

^{99m}Tc is a preferred imaging agent, mainly because it can be attached to an antibody at a position away from the antigen binding site. ^{99m}Tc may be linked to the antibody by a free Cys residue provided at or near the N- or C-terminus of the antibody. An antibody having such a free Cys residue may be prepared by recombinant DNA techniques, for

5   example by making a vector encoding the antibody-Cys fusion and expressing the antibody-Cys fusion in a host cell such as a bacterial host cell. The Cys residue is typically at the terminus of the antibody but may, for example, be the 2nd to 20th residue from the terminus.

An antibody according to the invention containing a detectable label is also useful

10   for radioimmunoguided surgery (RIGS, Blair et al., 1990; Tuttle et al., 1988) in addition to being useful for diagnosis of tumors. RIGS comprises administering a labeled antibody to a patient and thereafter surgically removing any tissue to which the antibody binds. Thus, the labeled antibody guides the surgeon towards tumor tissue and helps to distinguish from normal tissue.

15   Anti-tumor agents linked to an antibody may be any agent that destroys or damages a tumor to which the antibody has bound or in the environment of the cell to which the antibody has bound. For example, an anti-tumor agent may be a toxic agent such as a chemotherapeutic agent or a radioisotope. Suitable chemotherapeutic agents are known to those skilled in the art and include anthracyclines (e.g. daunomycin and doxorubicin),

20   methotrexate, vindesine, neocarzinostatin, cis-platinum, chlorambucil, cytosine arabinoside, 5-fluorouridine, melphalan, ricin and calicheamicin. The chemotherapeutic agents may be conjugated to the antibody using conventional methods (see e.g., Hermentin and Seiler, 1988).

Suitable radioisotopes for use as anti-tumor agents are also known to those skilled

25   in the art. For example, ^{131}I or ^{211}At may be used. These isotopes may be attached to the antibody using conventional techniques (see e.g., Pedley et al., 1993). The anti-tumor agent which is attached to the antibody may also be an enzyme which activates a prodrug. In this way, a prodrug may be administered which remains in its inactive form until it reaches the tumor site and is converted to its cytotoxin form. In practice, the antibody-enzyme

30   conjugate is administered to the patient and allowed to localize in the region of the tissue to

be treated. The prodrug is then administered to the patient so that conversion to the cytotoxic drug occurs in the region of the tissue to be treated. The anti-tumor agent conjugated to the antibody may also be a cytokine such as interleukin-2 (IL-2), interleukin-4 (IL-4) or tumor necrosis factor alpha (TNF-α). The antibody targets the cytokine to the tumor so that the cytokine mediates damage to or destruction of the tumor without affecting other tissues. The cytokine may be fused to the antibody at the DNA level using conventional recombinant DNA techniques.

It is understood that antibodies and antibody conjugates of the invention, where used in the human body for the purpose of diagnosis or treatment, will be administered in the form of a composition additionally comprising a pharmaceutically-acceptable carrier. Such carriers are well known to one of average skill in the art.

Throughout this application, various publications, patents, and patent applications have been referred to. The teachings and disclosures of these publications, patents, and patent applications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which the present invention pertains.

It is to be understood and expected that variations in the principles of invention herein disclosed may be made by one skilled in the art and it is intended that such modifications are to be included within the scope of the present invention.

The following examples further illustrate the invention.

Example 1

Materials and Methods

To test the MDD technique, a working model system was designed to isolate methyl-polymorphic DNA markers (see Fig. 1 for a schematic diagram of the MDD methods). There is evidence that in mammalian DNA, methylation patterns are both tissue-, and cell-type specific. Hence, DNA was isolated from two different types of cells from patients who have chronic lymphocytic leukemia (CLL). In CLL patients, the number of $CD5^+$ B type lymphocytes is abnormally high. These $CD5^+$ cells can be isolated by highly specific antibodies in a fluorescence activated cell sorter (FACs). For performing MDD, the DNA isolated from these $CD5^+$ cells served as the tester, while DNA isolated from neutrophils of the same patient served as the driver. A methyl-polymorphic DNA probe, named *CLL58*, which exhibited methyl-differential displays was isolated from an individual CLL patient.

After the successful isolation of these methyl-polymorphic probes, MDD was again successfully used in the isolation of methyl-polymorphic probes from breast cancer biopsies. Three methyl-polymorphic probes, *BR50*, *BR104*, and *BR254*, which gave methyl-differential displays in breast cancer and their matched normal cells, were isolated from three individual patients. *BR50* and *BR254* also gave methyl-differential displays in ovarian cancer and their matched normal cells, while the *BR254* probe exhibited a methyl-differential display in colon cancer cells.

DNA Isolation from malignant B-cells and neutrophil cells. Mononuclear and neutrophil cells were isolated from the peripheral whole blood of CLL patients by Ficoll/H-ypage centrifugation. Leukemic B-cells ($CD5^+$), were sorted from mononuclear cells in the presence of fluorescent anti-CD5 antibodies by FACs. DNAs were isolated by the phenol extraction method (Sambrook et al, 1989, pp. 9.16-9.19).

DNA from human cancer biopsies. Dissected human breast, ovarian, and colon cancer tissues (tumor and normal) were immediately frozen in liquid nitrogen, and stored at -70°C.

DNAs were isolated from both tumor and matched normal tissues by the phenol extraction method (Sambrook et al., 1989, pp. 9.16-9.19).

MDD. To prepare tester and driver amplicons, 1-2 μg of tester and driver genomic DNA were digested simultaneously with *Msp* I and *Mse* I restriction endonucleases (New England Biolabs). The digested DNAs were extracted with phenol/chloroform once, and precipitated with 2 volumes of ethanol after adding 1/10 volume of 3 M NaAc.

The DNAs were then resuspended in 10 μl of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH8.0). To ligate the first pair of oligonucleotide adapters (MSA24 with SEQ ID NO:1 and MSA12 with SEQ ID NO:2) to the digested tester and driver amplicon DNAs, 10 μl (1 μg) of digested DNA was mixed with: 2.5 μl of each of the unphosphorylated single strand adapters having SEQ ID NO:1 ($0.6\mu g/\mu l$) and SEQ ID NO:2 ($0.3\ \mu g/\mu l$); 2 μl of 10X ligation buffer (Boehringer Mannheim); and 3 μl of ddH$_2$0. The oligonucleotide adapters were annealed by cooling the mix gradually from 55°C to 10°C over a 1.5 hour period in a 4°C cold room, and then ligated to *Msp* I/*Mse* I-cut DNA fragments by overnight incubation with 1 U of T4 DNA ligase (Boehringer Mannheim) at 16°C. After ligation, the ligate was diluted with TEt buffer (10 mM Tris-HCl, 0.1 mM EDTA, 20 μg/ml of tRNA) at a concentration of 2 ng/μl.

To generate the tester and driver amplicons, tester and driver DNA ligated to the first set adapters were separately amplified by polymerase chain reaction ("PCR"). Two PCR tubes were set up to make the tester amplicon, while five PCR tubes were prepared for the driver amplicon. Each tube has a 400 μl reaction mixture containing 67 mM Tris-HCl, pH 8.8 at 25°C, 4 mM MgCl$_2$, 16 mM (NH$_4$)$_2$SO$_4$, 10 mM β-mercaptoethanol, 100 μg/ml of non-acetylated bovine serum albumin, 24 μl of dNTP (5 mM for each dATP, dCTP, dGTP, and dTTP), 8 μl of SEQ ID NO:1 (0.6 $\mu g/\mu l$), and 80 ng of DNA ligated to the adapters as templates. While the tubes were equilibrated at 72°C in a thermal cycler (Perkin-Elmer Cetus), 12U of Taq DNA polymerase (Perkin-Elmer Cetus), and 3U of pwo DNA polymerase (Boehringer Mannheim) were added. The reaction was overlaid with mineral oil and incubated for 5 minutes to fill-in the 3'-recessed ends of the ligates, and the DNAs were then amplified by 24 cycles of PCR. Each PCR cycle included 1 minute at

95°C, 1 minute at 67°C and 3 minutes at 72°C, and the last cycle was followed with a 10 minute extension time at 72°C. The DNA amplicons were extracted once with phenol, once with phenol/chloroform, and precipitated by 2 volumes of 100% ethanol after adding 1/4 volume of 10 M NH$_4$Ac. Amplicon DNAs were then resuspended in TE buffer.

To remove the first set of adapters from the driver amplicon, 80 μg of driver amplicon DNA was digested to completion with Msp I enzyme (10U/μg). The driver amplicon DNA was extracted once with phenol/chloroform, and after adding 1/4 volume of 10 M NH$_4$Ac, the DNA was precipitated with 2 volumes of 100% ethanol. The amplicon DNA was resuspended in TE buffer at a concentration of 0.5 μg/μl. To change the tester amplicon DNA adapters, 5 μg of tester amplicon was digested with Msp I (10U/μg). The digested tester amplicon DNA was purified by PCR purification spin column (Boehringer Mannheim).

For the hybridization/subtraction and amplification steps, Msp I digested tester amplicon DNA was ligated to a second set of adapters (MSB24 having SEQ ID NO:3 and MSB12 having SEQ ID NO:4): 1 μl (0.2 μg) of digested tester DNA was mixed with 1 μl of the SEQ ID NO:3 adaptor (0.6 μg/μl), 1 μl of the SEQ ID NO:4 adaptor (0.3 μg/μl), 1 μl of 10X ligation buffer, and 6 μl of ddH$_2$0. The same annealing and ligation steps were performed as for ligation of the first set of adapters. After ligation, the ligate was diluted with 40 μl of TEt buffer and mixed with 20 μg (40 μl) of driver amplicon DNA. The mix was extracted with phenol/chloroform once, and precipitated by adding 1/4 volume of 10 M NH$_4$Ac, and 2.5 volumes of 100% ethanol in a dry ice/ethanol bath for 10 minutes. The reaction solution was equilibrated to room temperature and the DNA was collected by centrifugation. The pellet was washed twice with 70% ice cold ethanol, dried, resuspended in 2 μl of 3XEE buffer. (Straus and Ausubel, 1990). After resuspension, the DNA was overlaid with one drop of mineral oil (Perkin-Elmer Cetus), denatured at 100°C for 5 minutes, and 0.5 μl of 5 M NaCl was added. The DNAs were hybridized at 68°C for at least 20 hours.

To amplify the difference products (DP1), the 3'-recessed ends were first filled in with DNA polymerase. The hybridized product was diluted with 198 μl of TE buffer, and a 50 μl of filling-in reaction was set up: 30 μl of diluted DNA hybrids, 10 μl of 5X PCR

buffer, 2 μl of dNTP (5 mM for each dATP, dCTP, dGTP, and dTTP), and 8 μl of ddH$_2$0. The reaction was equilibrated at 72°C, 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim) were added at 72°C for 10 minutes. The filled-in DNA hybrids were purified by a PCR purification spin column (Boehringer Mannheim).

5   To reduce the background for DP1 amplification, mung bean nuclease treatment of the filled-in DNA hybrids was performed: 9 μl of filled-in DNA hybrid solution mixed with 1 μl of 10X mung bean nuclease reaction buffer and 50U of mung bean nuclease (Boehringer Mannheim) was incubated at 30°C for 30 minutes. The reaction was terminated by heat inactivation at 100°C for 5 minutes after adding 40 μl of 50 mM Tris-HCl (pH 8.9).

10

To amplify DP1, 1 tube of 100 μl PCR reaction was set up. The reaction consisted of: 20 μl of mung bean nuclease-treated DNA hybrid solution, 20 μl of 5X PCR buffer, 5 μl of dNTP (5mM for each dATP, dCTP, dGTP. and dTTP), 2 μl of the SEQ ID NO:3 adaptor, and 52 μl of ddH$_2$0. The reaction was hot started at 94°C for 2 minutes after

15   adding 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim). The conditions for the PCR were 1 minute at 95°C, 1 minute at 67°C, and 3 minutes at 72°C for 35-40 cycles. The last cycle was followed with a 10 minutes extension step at 72°C. The DP1 was purified by PCR purification spin column (Boehringer Mannheim), and the concentration of the DNA was determined at

20   OD$_{260}$ in a spectrophotometer.

To prepare the second round of hybridization/subtraction and amplification steps, 3 μg of DP1 was digested with the restriction endonuclease *Msp* I (10U/μg) in a 30 μl reaction. The digested DP 1 was purified with PCR purification spin column (Beohringer

25   Mannheim).

To put a new set of adapters (MSC24 having SEQ ID NO:5 and MSC12 having SEQ ID NO:6) on DP1, 2 μl (0.1 μg) of DP1 was mixed with 1 μl of the MSC24 adaptor (0.6 μg/μl), 1 μl of the MSC12 adaptor (0.3 μg/μl), 1 μl of 10X ligation buffer (Beohringer Mannheim), and 6 μl of ddH$_2$0. The annealing and ligation reactions were

done as before. The ligate was diluted with 40 μl of TEt buffer to the concentration of 2 ng/μl.

To set up the hybridization reaction, 20 ng (10 μl) of diluted ligate was mixed with 20 μg (40 μl) of driver amplicon DNA. The DNA mix was extracted with an equal volume

5   of phenol/ chloroform once, and precipitated with 2.5 volumes of 100% ethanol after adding 1/4 volume of 10 M of NH$_4$Ac in a dry ice/ethanol bath for 10 minutes. The reaction solution was equilibrated to room temperature and the DNA was collected by centrifugation. The pellet was washed twice with 70% ice cold ethanol, dried, and resuspended in 2 μl of 3XEE. 0.5 μl of 5 M Nace was added after the DNA mix was

10   denatured at 100°C for 5 minutes. The hybridization reaction was carried out at 68°C for a minimum period of 20 hours.

To amplify the secondary different product (DP), the hybrids were diluted with 198 μl of TE buffer, and 100 μl of PCR amplification reaction was set up as follows: 4 μl of the diluted hybrid solution, 5 μl of dNTP (5 mM for each dATP, dCTP ,dGTP, dTTP), 20 μl

15   of 5X PCR reaction buffer, and 69 μl of ddH$_2$0. While the PCR tube was equilibrated at 72°C, 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim) were added to fill-in the 3'-recessed ends for 10 minutes. 2 μl of the MSC24 adaptor (0.6 μg/μl) was then added. The PCR reaction was hot started by raising the temperature to 94°C for 2 minutes, and 35 to 40 cycles of PCR

20   reaction were performed under the same conditions as were used for amplifying DP1.

DP usually contained several individual DNA fragments when electrophoresed on a 2% agarose gel. The individual DNA fragments were purified by DNA gel extraction kit (Qiagen Inc.), and subcloned into pUC118 vector, which was linearized by the restriction endonuclease *Acc* I. If the DP product had some DNA smear, it was treated with mung

25   bean nuclease as described above.

For each MDD, twelve to twenty colonies were chosen to amplify their inserts, from which different sized probes were selected for amplicon Southern blotting, and human genomic DNA Southern blotting.

Amplicon DNA Southern Blot. The first round of positive probe screening was performed with amplicon DNA Southern blots. Non-Radiation Southern blot and Detection Kits (Genius™) were purchased from Boehringer Mannheim. Probe labeling, and detection followed the instruction of the manufacturer.

Two to three μg of tester and driver amplicon DNA were electrophoresed on a 2% agarose gel, and blotted to positively charged nylon membranes (Boehringer Mannheim). For prehybridization, the membranes were placed at 68°C for 2-4 hours in solutions containing 6X SSC, 5X Denhardt's solution, 0.5% SDS, 0.1 M EDTA, and 50 μg/ml of salmon sperm DNA. Under the same conditions, the probes were added, and hybridized to the membranes overnight. The membranes were then rinsed three times with 2XSSC, 1X blot wash (12 mM Na_2HPO_4, 8 mM NaH_2PO_4, 1.4 mM $Na_4P_2O_7$, 0.5% SDS) at 68°C, and further washed three times (30 minutes for each time) with the same buffer at 68°C. Next, the membranes were equilibrated in buffer A (100 mM Tris-HCl, 150 mM, pH 7.5) and transferred into buffer B (2% black reagent in buffer A) solution and incubated at room temperature for one hour. The membranes were then washed 2 times (15 minutes for each time) with buffer A, and equilibrated in buffer C (100 mM Tris-HCl, 100 mM NaCl, 10 MM $MgCl_2$). Before the membranes were exposed on Kodak X-OMAT film, they were rinsed in lumi-P530 for 1 min and kept in a plastic sheet protector. Positive clones were identified by observing the probe hybridized more heavily to test amplicon DNA than to driver amplicon DNA.

Human Genomic DNA Southern Blot. The positive probes that were confirmed by the amplicon DNA Southern blot experiment were tested further by human genomic DNA Southern blotting. Genomic DNAs isolated from cancer tissues, and their respective matched normal tissues, were digested with *Msp* I restriction endonuclease, electrophoresed on 1.5% agarose gels, and transferred to Hybond Membranes (Amersham, Arlington Heights, IL). These membranes were exposed to UV light to immobilize the DNA. The probes for the Southern blot were labeled with High Prime DNA labeling kits (Boehringer Mannheim) following the instructions of the manufacturer. The

procedure for hybridization and blot wash were the same as in the Amplicon DNA Southern blotting section.

EXAMPLE 2

Different methylation patterns at CpCpG sites
in malignant B-cells and neutrophil cells

DNA isolated from malignant B-cells and neutrophil cells of four CLL patients (#58, #111, #112, #128) was used to perform MDD as described above in the Materials and Methods provided in Example 1. The DNA from leukemic B-cells was used as the tester, and the DNA from neutrophil cells was used as the driver. Tester and driver DNAs were simultaneously digested with the restriction endonucleases, *Msp* I and *Mse* I. After ligating the first set of adapters to the digested tester and driver DNA fragments, tester and driver amplicons were generated by PCR amplification (see Materials and Methods). After two rounds of hybridization/ subtraction and amplification, the difference products (DP) which were isolated from patient 111 appeared as individual DNA fragment bands on an agarose gel (Fig. 2). The DP fragments were then subcloned into the plasmid pUC118 (see Materials and Methods), and amplified by PCR. For the first round of selection, inserts of different sizes were selected and hybridized to the tester and driver amplicon DNAs from which they were derived in a non-radioactive Southern blot experiment (see materials and methods). The positive probes were selected if they hybridized a single band in the tester amplicon only, or if they hybridized to both tester and driver amplicon DNA, but the hybridization to the driver amplicon DNA band was far less intense. The final proof of a positive probe was that it displayed a different methylation pattern in a genomic DNA Southern blot experiment in which the parental malignant B-cell, and normal neutrophil cell genomic DNA, were digested with the restriction endonuclease *Msp* I (see materials and methods).

The *CLL58* probe isolated from patient #111 fulfilled these requirements. When *CLL58* was hybridized to its tester and driver amplicon DNA, the tester amplicon DNA gave a heavy hybridization band, while the driver amplicon DNA gave a much less intense hybridized band (Fig.3).

The *CLL58* probe was further tested by hybridization to a genomic Southern blot of *Msp* I-digested malignant B-cell and neutrophil cell DNA. The genomic Southern blot revealed that a larger fragment was much less intensely hybridized than a lower molecular weight fragment in the malignant B-cell DNA, while in the neutrophil cell DNA, just the opposite result occurred (Fig. 4). These results demonstrate that in human genomic DNA, there are methylated external cytosine residues at CpCpG sequences, and those CpCpG sequences are differentially methylated in different types of cells. Moreover, the *CLL58* probe can detect and distinguish cell-type specific methylation patterns.

EXAMPLE 3

MDD-Isolated Probes from Human Breast Cancer Tumors
Also Detect Ovarian and Colon Cancers

We have proven the existence of methylation of the external cytosine residue at CpCpG sites in the human genome, and have isolated CpCpG site related methyl-polymorphic markers in a working model system. MDD was then applied to successfully isolate CpCpG related methyl-polymorphic markers from human breast cancer biopsies.

Isolation of Probes From Breast Tumor DNA

MDD was used to test five pairs of DNA samples (tumor DNA and matched normal DNA) obtained from five breast cancer patients. To isolate methyl-polymorphic markers, tumor DNA served as the tester, and normal DNA served as the driver. MDD was performed in the same manner as described in Example 1. Individual DNA fragments were isolated by MDD from three breast cancer patients after two rounds of hybridization/subtraction and amplification (Fig. 5).

Three probes were further analyzed, *BR50*, *BR104*, and *BR254*, isolated from patient #14, patient #13, and patient #4, respectively. Each of these three probes appropriately hybridized with their own tester amplicon DNA, but also hybridized slightly with the driver amplicon DNA (Fig. 6a-c).

These probes were then examined further by hybridizing them to *Msp* I-digested parental tumor and matched normal genomic DNAs. The genomic Southern blot for probe

BR50 showed that the probe hybridized with a lower band to a much greater degree than to an upper band in tumor DNA (Fig. 7a). However, in normal DNA, just the opposite occurred -- the *BR50* probe hybridized to the upper band more than to the lower band (Fig. 7a).

The genomic DNA Southern blot for probe *BR104* showed that the probe hybridized with a lower and a middle band in the tumor DNA, while in the normal DNA, probe *BR104* hybridized with a middle and an upper band (Fig. 7b).

For the probe *BR254*, the Southern blot result showed that it hybridized with a lower band and a upper band in the tumor DNA, however, it only hybridized with an upper band in the normal DNA (Fig. 7c).

The results of three cases, indicate that in normal cells the genomic DNA is more highly methylated, is cut less frequently and gives rise to larger fragments than is observed for DNA from tumor cells. The isolated probes therefore hybridize to higher molecular weight bands in normal DNA. In contrast, the lower bands which were hybridized with probes *BR50*, *BR104*, and *BR254*, are generated by demethylation of the external C residue at CpCpG DNA sites in the tumor cells. The weakly hybridized upper bands in the tumor DNA are likely caused by normal cell contamination.

Breast Tumor Probes Detect Hypomethylation in Other Cancer Patients

To understand whether this hypomethylation event also happens in other breast cancer patients, normal cell and tumor cell DNA was isolated from ten breast cancer patients, cleaved by the *Msp* I enzyme, and probed with *BR50* and *BR104*. Of the ten breast patients whose DNA was probed with *BR50*, five had a hybridization pattern which was similar to that of the patient from which *BR50* was isolated (the "parental patient") (Fig. 8). Of the ten breast patients whose DNA was probed with *BR104*, five had a hybridization pattern similar to that of the parental patient. (See Fig. 9, providing four of these five patterns).

The *BR254* probe was hybridized to normal and tumor DNA of 16 breast cancer patients. Of the 16 patients, 10 had a hybridization pattern which was similar to that of the parental patient. However, some differences in the methylation patterns of these patients

were observed. Some patients only had the lower band in their tumor DNA, and the upper band in their normal DNA (Fig. 10). However, other patients had the lower band in their tumor DNA, but had both lower and upper bands in their normal DNA (Fig. 10). These results indicate that the external cytosine residue at CpCpG sites are fully demethylated in some breast tumors. However, the lower bands appearing in normal DNAs probably were caused by a partial demethylation event, which suggests that the normal cells were under the process of neoplastic change. In general, the different extent of demethylation identified by probes *BR50*, *BR104*, and *BR254* in human breast cancer patients is a common phenomena and may be used to predict the extent of neoplastic change.

Detection of a DNA Amplification Event in a Breast Cancer Patient. The *BR254* probe detected a highly amplified DNA event in one of the 16 breast cancer patients tested with that probe (Fig. 11). This result provides evidence that imbalanced DNA methylation may cause a conformational change of chromatin which then induces a DNA amplification.

Hypomethylation Detection by *BR50, BR104*, and *BR254* in Ovarian Cancers. To examine whether the same hypomethylation phenomena also happens in human ovarian cancer, DNA samples isolated from 8 patients with ovarian cancer were analyzed. For some patients, two DNA samples were obtained, one from primary ovarian cancer tissue, and the other from their matched normal tissue. For other patients, three DNA samples were obtained, they were from primary ovarian cancer tissue, metastatic tumor tissue, and matched normal tissue. The DNAs were cleaved with *Msp* I, and hybridized with the probes *BR50* and *BR104* in Southern blot experiments.

Of the eight ovarian patients tested with the *BR50* probe, five patients had similar hybridization patterns: the *BR50* probe hybridized almost equally to lower and upper bands in their normal DNA samples, while in the primary tumor DNA, the upper band was only slightly hybridized. However, in the metastatic tumor DNA of these 5 patients, only the lower band was hybridized (Fig. 12).

No different hybridization patterns were detected using the probe *BR104*. This result indicates that the *BR104* probe was detecting a tissue-specific hypomethylation event.

Eleven patients' DNA samples were examined using the *BR254* probe. The Southern blot results showed that of 11 patients, 4 had the same hybridization patterns which appeared in the parental DNA pair (Fig. 13).

These results suggest that probes *BR50* and *BR254* detect tumor-related hypomethylation events in both breast and ovarian cancer cells. Interestingly, in the case of probe *BR50*, the complete disappearance of the upper band in the metastatic tumor DNA indicates that the extent of demethylation seems to be related to the progression of the disease.

Hypomethylation Detection using *BR50*, *BR104*, and *BR254* on Colon Cancer DNA. The methylation patterns of tumor and matched normal DNA samples isolated from 10 colon cancer patients were examined using *BR50*, *BR104*, and *BR254* probes. DNAs were digested by the *Msp* I enzyme, Southern blots prepared and the blots were hybridized with these probes.

No difference in the methylation patterns of colon tumor DNA and normal DNA were detected by probes *BR50* and *BR104* (data not shown), indicating that both probes were detecting a tissue-specific hypomethylation event in colon cells.

However, for the probe *BR254*, different methylation patterns were observed. Of the ten colon cancer patients tested, six had hypomethylation patterns opposite to those observed for the breast and ovarian cancer patients. Instead of hybridizing more to an upper band in the normal DNA samples, probe *BR254* hybridized with the lower band. Moreover, instead of hybridizing more to the lower band in the tumor DNA, *BR254* hybridized equally with the lower band and a upper band (Fig. 14). These results demonstrate that while the *BR254* probe detects hypomethylation in breast cancer cells, in colon cancer DNA the *BR254* probe detects hypermethylation. Therefore, methylation patterns are not only tissue-type specific, but cancer-specific, and can be different in different types of cancers.

Sequencing and Chromosome Assignment of Probes *BR50* and *BR254*. DNA sequencing showed that *BR50* contains 1005 bp, with a CG content of 58% (SEQ ID

NO:10); and *BR254* contains 332 bp, with a CG content of 55% (SEQ ID NO:8). Both probes have the nucleotide sequence GGCCGG at one end, and the nucleotide sequence CCGG sequence at the other end. *Msp* I, which cleaves at CCGG sequences, is sensitive to methylation at the external C residue, and will not cleave CmCGG. Thus, CCGG sequences at the ends of both probes were methylated in normal breast DNA and hypomethylated in tumor DNA.

Probes *BR50* and *BR254* were assigned to chromosome #3 and #8 respectively by genomic Southern blotting of the *Hin*d III digested monochromosonal human/rodent somatic cell hybrid mapping panel #2 (NIGMS Human Genetic Mutant Cell Repository). showed that they are located on chromosome #3, and #8, respectively. Fine chromosome mapping was performed using GeneBridge 4 Radiation Hybrid Panel (Research Genetics, Inc.) as templates (Walter et al., 1994). *BR50* is proximally located on 3p12-14, and *BR254* is proximally located on 8q11-12. The adjacent Sequence Tagged Site (STS) to *BR50* is AFMB362WB9, and to *BR254* is WI-3862. A known oncogene, Fibroblast Growth Factor Receptor (FGFR) has been located on 8q11-12 region, and our results did show that the genomic region corresponding to *BR254* was amplified in a breast cancer biopsy.

Determination of Coding Sequences. 2×10^6 plaques from a human placenta genomic phage library, EMBL3 SP6/7 (ClonTech, Inc.) were screened with the probe *BR50*. A phage clone hybridizing to *BR50* was obtained. The cloned genomic DNA was released by digestion with *Sst* I and transferred into pUC118. 8 DNA fragments were generated, and the total length of these fragments was approximately 17 kbp. Plasmids containing fragments smaller than 1 kbp were completely sequenced, while plasmids containing fragments larger than 1 kbp were partially sequenced. Homology searches against NIH GenBank revealed two exons, designated BR50-44 and BR50-45, consisting respectively of 112 and 132 nucleotides. Both regions displayed homology to several mammalian proteases, such as serine proteases and tryptases. Based on the sequence information obtained, oligomers were designed for the purpose of amplifying a cDNA sequence from a human universal cDNA library panel (ClonTech Inc.) A PCR product of about 682 bp

was generated using one oligomer (SEQ ID NO:11) based on the sequence obtained from BR50-44 and a second oligomer (SEQ ID NO:12) based on the sequence obtained from BR50-45. This PCR product was directly sequenced. We obtained an incomplete coding sequence of 854bp by combining the sequence obtained for the PCR product with

5   sequences from BR50-44 and BR50-45. A DNA homology search against NIH GenBank again revealed similarity to serine proteases.

Tissue specific expression. Two commercially available tissue specific northern blot panels (Human Multiple Tissue Northern Blot, Human Multiple Tissue Northern Blot II,

10   Clontech Inc.) containing poly A RNA from a total of 16 different human tissues were tested for the expression of an mRNA capable of hybridizing to the coding sequence described above. A probe (SEQ ID NO:13) corresponding to the portion of the coding sequence amplified by the oligomers of SEQ ID NO:11 and SEQ ID NO:12 hybridized strongly and specifically to a 1.3kb testes specific transcript. Hybridization to the remaining

15   15 human tissues that were tested was not detected (Fig. 15). Therefore the locus was designated TSP50 (Testes Specific Protease).

The methylation status of TSP50 was investigated in nine normal human tissues. 6 μg of DNA from each tissue was cleaved with *Hpa* II (Fig. 6a) or *Msp* I (Fig. 6b) and analyzed by Southern hybridization with probe *BR50*. Among DNAs digested with *Hpa* II,

20   hybridized bands of approximately 1 kbp and 2 kbp are observed only in digested DNA from testes tissue. TSP50 DNA from other tissues are sufficiently methylated that digestion products of 1 kb or 2 kb are undetectable. Like *Msp* I, *Hpa* II recognizes and cleaves DNA at nonmethylated CCGG sequences. Unlike *Msp* I which will not cleave a CCGG sequence when the outer cytosine is methylated, *Hpa* II will not cleave a CCGG sequence

25   when either cytosine residue is methylated. Among DNAs digested with *Msp* I, one or both fragments are observed in all samples. Where a population of CCGG sequences is not fully methylated or demethylated, digestion of DNA samples with *Msp* I will yield the type of results seen in Fig. 6b where intermediate degrees of digestion are observed. Notably, the predominant TSP50 DNA fragment from testes is the 1 kb fragment. Less of the 2 kb

30   fragment is present than in any other sample. This indicates that TSP50 DNA from testes is

less methylated than TSP50 DNA from all other tissues tested. In both experiments of Fig. 6, hybridization of a control probe to a single band in each digested DNA sample indicates that digestion was complete.

Levels of TSP50 mRNA were investigated in breast cancer tissues from eighteen subjects as compared to matched normal breast tissue. TSP50 mRNA was detected by RT-PCR. PCR products corresponding to TSP50 were generated from mRNA from five cancer tissues, but not from normal tissue from the same subjects (Fig. 17). This was consistent with preliminary results which indicated that TSP50 was differentially methylated in approximately 40% of breast cancer tissues tested.

Isolation of full length of cDNA. To obtain a full length TSP50 cDNA, a human testes cDNA library (Clontech, Inc.) was screened with the *TSP50* probe having the nucleotide sequence given by SEQ ID NO:13. A cDNA clone consisting of 1,240 bp was obtained (SEQ ID NO:14). Sequence analysis indicates that this fragment encodes a protein containing 385 amino acids (SEQ ID NO:16). A stop codon located 117 bp upstream of the putative translation initiation site, and a 125 bp 3' untranslated region before a poly adenosine sequence suggest that a full length coding sequence has been obtained.

References

Aujame, L., Geoffroy, F. and Sodoyer, R. (1997) High affinity human antibodies by phage display. Hum. Antibodies 8:155-168.

Barbas, C.F. 3d, Kang, A.S., Lerner, R.A. and Benkovic, S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proc. Natl. Acad. Sci. USA 88:7978-7982.

Blair, S.D., Theodorou, N.A., Begent, R.H., Dawson, P.M., Salmon, M., Riggs, S., Kelly, A., Boxer, G., Southall, P. and Gregory, P. (1990) Comparison of anti-fetal colonic microvillus and anti-CEA antibodies in peroperative radioimmunolocalisation of colorectal cancer. Br. J. Cancer 61:891-894.

Brüggemann, M. and Taussig, M.J. (1997) Production of human antibody repertoires in transgenic mice. Curr. Opin. Biotechnol. 8:455-458.

Burton, D.R. and Barbas, C.F. 3rd (1994) Human antibodies from combinatorial libraries. Adv. Immunol. 57:191-280.

Clark, S.J., Harrison, J. and Frommer, M. (1995) CpNpG methylation in mammalian cells. Nat. Genet. 10:20-27.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1991) Making antibody fragments using phage display libraries. Nature 352:624-628.

Elliott, M.J., Maini, R.N., Feldmann, M., Long-Fox, A., Charles, P., Bijl, H. and Woody, J.N. (1994) Repeated therapy with monoclonal antibody to tumour necrosis factor alpha (cA2) in patients with rheumatoid arthritis. Lancet 344:1125-1127.

Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinson, I.M., Waterhouse, P., Crosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., et al. (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. EMBO J. 13:3245-3260

Goldenberg, D.M. (1992) Cancer imaging with CEA antibodies: historical and current perspectives. Int. J. Biol. Markers 7:183-188.

Green, L.L., Hardy, M.C., Maynard-Currie, C.E., Tsuda, H., Louie, D.M., Mendez, M.J., Abderrahim, H., Noguchi, M., Smith, D.H., Zeng, Y., et al. (1994) Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. Nat. Genet. 7:13-21.

Hawkins, R.E., Russell, S.J, and Winter, G. (1992) Selection of phage antibodies by binding affinity. Mimicking affinity maturation. J. Mol. Biol. 226:889-896.

Henry, I., Bonaiti-Pellie, C., Chehensse, V., Beldjord, C., Schwartz, C., Utermann, G. and Junien, C. (1991) Uniparental paternal disomy in a genetic cancer-predisposing syndrome. Nature 351:665-667.

Hermentin, P. and Seiler, F.R. (1988) Investigations with monoclonal antibody drug (anthracycline) conjugates. Behring Inst. Mitt. 82:197-215.

Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J. and Lerner, R.A. (1989) Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science 246:1275-1281.

Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S. and Winter, G. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature 321:522-525.

Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. and Foeller, C. (1991) Sequences of Proteins of Immunological Interest. 5th edition. 3 vols. Bethesda, MD: National Institutes of Health. National Center for Biotechnology Information. NIH publication no. 91-3242

Kwoh, D.Y. and Kwoh, T.J. (1990) Target amplification systems in nucleic acid-based diagnostic approaches. Am. Biotechnol. Lab. 8:14-25.

Laird, P.W. and Jaenisch, R. (1994) DNA methylation and cancer. Hum. Mol. Genet. 3:1487-1495.

Lisitsyn, N., Lisitsyn, N. and Wigler, M. (1993) Cloning the differences between two complex genomes. Science 259:946-951.

Little, M. and Wainwright, B. (1995) Methylation and p16: suppressing the suppressor. Nat. Med. 1:633-634.

LoBuglio, A.F., Wheeler, R.H., Trang, J., Haynes, A., Rogers, K., Harvey, E.B., Sun, L., Ghrayeb, J. and Khazaeli, M.B. (1989) Mouse/human chimeric monoclonal antibody in man: kinetics and immune response. Proc. Natl. Acad. Sci. USA 86:4220-4224.

Low, N.M., Holliger, P.H. and Winter, G. (1996) Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. J. Mol. Biol. 260:359-368

McCafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348:552-554.

Nicholls, R.D., Knoll, J.H., Butler, M.G., Karam, S. and Lalande, M. (1989) Genetic imprinting suggested by maternal heterodisomy in nondeletion Prader-Willi syndrome. Nature 342:281-285.

Pedone, P.V., Tirabosco, R., Cavazzana, A.O., Ungaro, P., Basso, G., Luksch, R., Carli, M., Bruni, C.B., Frunzio, R. and Riccio, A. (1994) Mono- and bi-allelic expression of insulin-like growth factor II gene in human muscle tumors. Hum. Mol. Genet. 3:1117-1121.

Ogawa, O., Eccles, M.R., Szeto, J., McNoe, L.A., Yun, K., Maw, M.A., Smith, P.J. and Reeve, A.E. (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumor. Nature 362:749-751.

Padlan, E.A. (1991) A possible procedure for reducing the immunogenicity of antibody variable domains while preserving their ligand-binding properties. Mol. Immunol. 28:489-498.

Pedley, R.B., Boden, J.A., Boden, R., Dale, R. and Begent, R.H. (1993) Comparative radioimmunotherapy using intact or F(ab')2 fragments of 131I anti-CEA antibody in a colonic xenograft model. Br. J. Cancer 68:69-73.

Queen, C., Schneider, W.P., Selick, H.E., Payne, P.W., Landolfi, N.F., Duncan, J.F., Avdalovic, N.M., Levitt, M., Junghans, R.P. and Waldmann, T.A. (1989) A humanized antibody that binds to the interleukin 2 receptor. Proc. Natl. Acad. Sci. USA 86:10029-10033.

Rainier, S., Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E. and Feinberg, A.P. (1993) Relaxation of imprinted genes in human cancer. Nature 362:747-749.

Rainier, S., Dobry, C.J. and Feinberg, A.P. (1995) Loss of imprinting in hepatoblastoma. Cancer Res. 55:1836-1838.

Riechmann, L., Clark, M., Waldmann, H. and Winter, G. (1988) Reshaping human antibodies for therapy. Nature 332:323-327.

Roguska, M.A., Pedersen, J.T., Keddy, C.A., Henry, A.H., Searle, S.J., Lambert, J.M., Goldmacher, V.S., Blattler, W.A., Rees, A.R. and Guild, B.C. (1994) Humanization of murine Mabs through variable domain resurfacing. Proc. Natl. Acad. Sci. USA 91:969-973.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487-491.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Vols. 1-3 (Cold Spring Harbor Press, Cold Spring Harbor, NY).

Straus, D. and Ausubel, F.M. (1990) Genomic subtraction for cloning DNA corresponding to deletion mutations. Proc. Natl. Acad. Sci. USA 87:1889-1893. Published erratum appears in Proc. Natl. Acad. Sci. USA 88:1590.

Suzuki, H., Veda, R., Takahashi, T. and Takahashi, T. (1994) Altered imprinting in lung cancer. Nat. Genet. 6:332-333.

Tuttle, S.E., Jewell, S.D., Mojzisik, C.M., Hinkle, G.H., Colcher, D., Schlom, J. and Martin, E.W. (1988) Intraoperative radioimmunolocalization of colorectal carcinoma with a hand-held gamma probe and MAb B72.3: comparison of in vivo gamma probe counts with in vitro MAb radiolocalization. Int. J. Cancer 42:352-358.

Vaughan, T.J., Osbourn, J.K. and Tempest, P.R. (1998) Human antibodies by design. Nat. Biotechnol. 16:535-539

Waldmann, T.A. (1991) Monoclonal antibodies in diagnosis and therapy. Science 252:1657-1661.

Walter, M.A., Spillett, D.J., Thomas, P., Weissenbach, J. and Goodfellow, P.N. (1994) A method for constructing radiation hybrid maps of whole genomes. Nat. Genet. 7:22-28.

Wagner, S.D., Popov, A.V., Davies, S.L., Xian, J., Neuberger, M.S. and Brüggemann, M. (1994) The diversity of antigen-specific monoclonal antibodies from transgenic mice bearing human immunoglobulin gene miniloci. Eur. J. Immunol. 24:2672-2681

Willison, K. (1991) Opposite imprinting of the mouse Igf2 and Igf2r genes. Trends Genet. 7:107-109.

Yang, W.P., Green, K., Pinz-Sweeney, S., Briones, A.T., Burton, D.R. and Barbas, C.F. 3rd (1995) CDR walking mutagenesis for the affinity maturation of a potent human anti-HIV-1 antibody into the picomolar range. J. Mol. Biol. 254:392-403.

Zhan, S., Shapiro, D.N. and Helman, L.J. (1994) Activation of an imprinted allele of the insulin-like growth factor II gene implicated in rhabdomyosarcoma. J. Clin. Invest. 94:445-448.